Detection of Polymorphisms in the Human 5-Lipoxygenase Gene

Related Applications

This application claims priority to U.S. Provisional Application No. 60/267,515 filed on February 8, 2001 and Application No.60/314,248 filed August 21, 2001, the contents of which are incorporated herein by reference.

Background of the Invention

5-lipoxygenase (5-LO) is the first committed enzyme in the pathway leading to leukotriene synthesis and is responsible for the conversion of arachidonic acid to LTA4 via the unstable intermediate 5-hydroperoxy eicosotetraenoic acid (Samuelsson *et al.*, (1987) *Science* 237:1171; Samuelsson (1983) *Science* 220:568). Leukotriene A4 can, in turn, be converted to leukotriene B4, C4, D4, or E4. Leukotrienes are potent local mediators which influence inflammatory and allergic response, including asthma, rheumatoid arthritis, psoriasis, thrombotic disease, ulcerative colitis, bronchitis, sinusitis, allergic and non-allergic rhinitis, and lupus. Leukotrienes (LTs) B4, C4, D4, and E4 have been shown experimentally to mimic the pathologic changes seen in asthma and to play a role in several inflammatory mechanisms that lead to airflow obstruction in asthma. Such mechanisms include bronchoconstriction, mucosal edema, increased secretion of mucus, and an inflammatory-cell infiltrate that is rich in eosinophils. (O'Bryne (1997) *Chest* 111:27S-34S). The slow-reacting substance of anaphylaxis is now known to be a mixture of leukotrienes C4, D4, and E4, all of which are potent bronchoconstrictors that exert their biological actions through specific ligand-receptor interactions thereby increasing vascular permeability and constricting smooth muscle (U.S. Patent Number 5,750,565; Silverman, *et al.* (1998) *Clin Exp Allergy* 28 Suppl 5:164-70).

There have been research efforts to develop specific receptor antagonists or inhibitors of leukotriene biosynthesis, to prevent or minimize pathogenic inflammatory responses mediated by these compounds, including the inhibition of 5-LO. For example, European Patent Application Nos. 90117171.0 and 901170171.0 disclose indole, benzofuran, and benzothiophene lipoxygenase inhibiting compounds (U.S. Patent Number 5,750,565). It has been established that treating patients with agents that have the capacity to inhibit 5-LO results in improvement in lung function, reduction in asthma symptoms, and decreased need

for alternative asthma treatments (Persson *et al.*, (1995) *Anesthesiology* 82:969; Israel *et al.*, (1993) *Ann. Int. Med.* 119:1059). However, in patients with asthma, there is a heterogeneous response to treatment with 5-LO inhibitors. Mutations in the 5-LO gene sequence may be associated with responsiveness to therapy and/or susceptibility to diseases or disorders, *e.g.*, asthma.

The 5-LO gene has been cloned, both as a cDNA (Matsumoto *et al.*, (1988) *Proc. Natl. Acad. Sci. USA* 85:3406; Dixon *et al.*, (1988) *Proc. Natl. Acad. Sci. USA* 85:416; Balcarek *et al.*, (1988) *J. Biol. Chem.* 263:13937) and as a genomic clone (Hoshiko *et al.*, (1990) *Proc. Natl. Acad. Sci. USA* 87:9073; Funk *et al.*, (1989) *Proc. Natl. Acad. Sci. USA* 86:2587). The 5-LO gene is approximately 85 kilobases in size and contains 14 exons and 15 introns. The region 88 to 212 base pairs upstream of the 5-LO translation start site contains a number of sequences known to be recognition sites for transcriptional regulators (Hoshiko *et al.*, (1990) *Proc. Natl. Acad. Sci. USA* 87:9073, incorporated herein by reference). For example, binding sites for AP1 and Sp1, each of which can act as either a transcriptional activator or a transcriptional repressor, depending on context, are found in this region. With respect to the Sp1 binding site, it comprises 5 tandem Sp1 motifs (GGGCGG) found -147 to -176 base pairs upstream of the ATG translation start site. Deletions of the Sp1 motifs reduce transcription from this 5' upstream regulatory element (Hoshiko *et al.* (1990) *Proc. Natl. Acad. Sci. USA* 87:9073). Sp1 binding sites comprising such motifs are similarly found in the promoter region of a variety of other genes (Li *et al.*, (1995) *Gene* 164:229; Wariishi *et al.*, (1995) *Biochem. Biophys. Res. Commun.* 216:729; Tang *et al.*, (1995) *Biochem. Biophys. Res. Commun.* 213:673; Khachigian *et al.*, (1995) *J. Biol. Chem.* 270:27679).

Certain polymorphisms in the 5-LO gene have been shown to be correlated with patient responsiveness to 5-LO inhibitor therapy (U.S. Patent Number 6,090,547 by Drazen, *et al.* (2000)). Drazen *et al.* disclosed that asthma patients having polymorphisms associated with reduced 5-LO gene expression are less responsive to 5-LO inhibitor therapy than patients having normal 5-LO gene expression. Specifically, Drazen *et al.* showed that 5-LO promoters comprising a variant Sp1 binding site having 3, 4 or 6 Sp1 motifs were less active than the wild-type 5-LO promoter comprising a Sp1 bind site having 5 Sp1 motifs, and that

asthma patients homozygous or heterozygous for 5-LO alleles comprising 5-LO promoters having such variant Sp1 binding sites were less responsive to 5-LO inhibitor therapy than patient homozygous for the wild-type 5-LO allele. Drazen *et al.* also showed that the 3 Sp1 motif binding site is in linkage disequilibrium with a G to A polymorphism in exon 2 of the 5-LO gene, and that the 4 Sp1motif binding site is in linkage disequilibrium with a C to T polymorphism in exon 1of the 5-LO gene.

It would also be desirable to identify additional polymorphisms within the 5-LO gene which are in linkage disequilibrium with a variant Sp1 binding site in the 5-LO promoter. It would further be desirable to provide prognostic, diagnostic, pharmacogenomic and therapeutic methods utilizing the identified polymorphisms.

Summary of the Invention

The present invention relates to polymorphisms in the 5-LO gene. The invention is based, in part, on the discovery of a novel 5-LO haplotype which comprises five sequence polymorphisms within the 5-LO gene that are in linkage disequilibrium with each other. The five 5-LO polymorphisms are set forth in Table 1. Three of these, 5loprr1, 5lo01a and 5lo04a, are newly identified polymorphisms. The other two, 5lonrra and 5lonrrb, have been described by In *et al.* in *J. Clinical Invest.* 99:1130-1137 (1997). The present invention is based, also in part, on the discovery that the haplotype of the invention is in linkage disequilibrium with any one of three variant Sp1 binding sites in the 5-LO promoter region that were previously reported by Drazen *et al.* (U.S. Patent No. 6,090,547). Accordingly, any one of the five polymorphisms of the invention is a marker for the other four polymorphisms and a variant Sp1 binding site having 3, 4 or 6 Sp1 motifs in the 5-LO promoter region. Moreover, since Drazen *et al. (id.)* demonstrated that asthma patients homozygous or heterozygous for a 5-LO allele comprising any one of such variant Sp1 binding sites are less susceptible to effective treatment with 5-LO inhibitor therapy, the haplotype of the invention is also a marker of reduced responsiveness amongst inflammatory or allergy disease patients to 5-LO inhibitor therapy.

The present invention is based further, in part, on the discovery that the 5lo01a polymorphism of the invention is associated with an abnormally low eosinophil count. That

is, individuals heterozygous or homozygous for the 5lo01a polymorphism have been found to have significantly lower eosinophil levels compared to individuals not having the polymorphism. Thus, the haplotype of the invention and each of its component polymorphisms are all markers of low eosinophil levels.

The invention provides a method for identifying an inflammatory or allergy disease patient who is less susceptible to effective treatment with a 5-LO inhibitor, comprising determining the presence or absence of a 5-LO allelic variant comprising a sequence selected from the group consisting of those set forth in SEQ ID NO: 4, SEQ ID NO:5, and SEQ ID NO:6, or the complement of the sequence in a nucleic acid sample from an inflammatory or allergy disease patient, wherein the presence of the 5-LO allelic variant in the sample indicates that the patient is less susceptible to effective treatment with a 5-LO inhibitor. In one embodiment, the 5-LO allelic variant comprises the sequences of those set forth in SEQ ID NO:4, SEQ ID NO:5, and SEQ ID NO:6, or the complement of each of those sequences. In another embodiment, the 5-LO allelic variant comprises the sequences of those set forth in SEQ ID NO:4, SEQ ID NO:5, and SEQ ID NO:6, SEQ ID NO:7, and SEQ ID NO:8, or the complement of each of those sequences. In yet another embodiment, the 5-LO allelic variant comprises: (a) a sequence selected from the group consisting of those set forth in SEQ ID NO: 4, SEQ ID NO:5, and SEQ ID NO:6, or the complement of each of those sequences; and (b) a variant Sp1 binding site having a 3, 4 or 6 Sp1 motif. In yet another embodiment, the 5-LO allelic variant comprises: (a) a sequence selected from the group consisting of those set forth in SEQ ID NO: 4, SEQ ID NO:5, and SEQ ID NO:6, SEQ ID NO:7, and SEQ ID NO:8, or the complement of each of those sequences; and (b) a variant Sp1 binding site having a 3, 4 or 6 Sp1 motif.

The invention also provides a method for identifying a patient with an abnormally low eosinophil level, comprising determining the presence or absence of a 5-LO allelic variant comprising a sequence selected from the group consisting of those set forth in SEQ ID NO: 4, SEQ ID NO:5, and SEQ ID NO:6, or the complement of the sequence in a nucleic acid sample from a patient, wherein the presence of the 5-LO allelic variant in the sample indicates that the patient has an abnormally low eosinophil level. In one embodiment, the 5-LO allelic variant comprises the sequences of those set forth in SEQ ID NO:4, SEQ ID

NO:5, and SEQ ID NO:6, or the complement of each of those sequences. In another embodiment, the 5-LO allelic variant comprises the sequences of those set forth in SEQ ID NO:4, SEQ ID NO:5, and SEQ ID NO:6, SEQ ID NO:7, and SEQ ID NO:8, or the complement of each of those sequences. In yet another embodiment, the 5-LO allelic variant comprises: (a) a sequence selected from the group consisting of those set forth in SEQ ID NO: 4, SEQ ID NO:5, and SEQ ID NO:6, or the complement of each of those sequences; and (b) a variant Sp1 binding site having a 3, 4 or 6 Sp1 motif. In yet another embodiment, the 5-LO allelic variant comprises: (a) a sequence selected from the group consisting of those set forth in SEQ ID NO: 4, SEQ ID NO:5, and SEQ ID NO:6, SEQ ID NO:7, and SEQ ID NO:8, or the complement of each of those sequences; and (b) a variant Sp1 binding site having a 3, 4 or 6 Sp1 motif. In a preferred embodiment, the patient is an inflammatory or allergy disease patient. In a more preferred embodiment, the patient is an asthma patient.

It is believed that inflammatory or allergy disease patients having abnormally low eosinophil levels have milder disease symptoms than patients having normal eosinophil levels. Accordingly, the instant invention further provides a method for determining the severity of disease symptoms in an inflammatory or allergy disease patient comprising determining the presence or absence of a 5-LO allelic variant comprising a sequence selected from the group consisting of those set forth in SEQ ID NO: 4, SEQ ID NO:5, and SEQ ID NO:6, or the complement of the sequence in a nucleic acid sample from the patient, wherein the presence of the 5-LO allelic variant in the sample indicates that the patient has or will less severe disease symptoms than patients having wild-type 5-LO alleles. In one embodiment, the 5-LO allelic variant comprises the sequences of those set forth in SEQ ID NO:4, SEQ ID NO:5, and SEQ ID NO:6, or the complement of each of those sequences. In another embodiment, the 5-LO allelic variant comprises the sequences of those set forth in SEQ ID NO:4, SEQ ID NO:5, and SEQ ID NO:6, SEQ ID NO:7, and SEQ ID NO:8, or the complement of each of those sequences. In yet another embodiment, the 5-LO allelic variant comprises: (a) a sequence selected from the group consisting of those set forth in SEQ ID NO: 4, SEQ ID NO:5, and SEQ ID NO:6, or the complement of each of those sequences; and (b) a variant Sp1 binding site having a 3, 4 or 6 Sp1 motif. In yet another

embodiment, the 5-LO allelic variant comprises: (a) a sequence selected from the group consisting of those set forth in SEQ ID NO: 4, SEQ ID NO:5, and SEQ ID NO:6, SEQ ID NO:7, and SEQ ID NO:8, or the complement of each of those sequences; and (b) a variant Sp1 binding site having a 3, 4 or 6 Sp1 motif. In a preferred embodiment, the patient is an asthma patient.

The invention also provides isolated nucleic acids comprising the novel 5-LO polymorphisms and the haplotype of the invention. The nucleic acid molecules of the invention include specific 5-LO allelic variants which differ from the reference 5-LO sequences set forth in SEQ ID NO:1 (GI 187166), SEQ ID NO: 2 (GI 8247778), SEQ ID NO:3 (GI 4502056). The preferred nucleic acid molecules of the invention comprise newly identified 5-LO allelic variants or portions thereof having the any of the novel polymorphisms shown in Table 1 (*i.e.*, those comprising the sequence of any of those set forth in SEQ ID NOs:4-6 or the complement thereof), polymorphisms in linkage disequilibrium with the polymorphisms shown in Table 1, and combinations thereof. The haplotype also includes known polymorphisms. These known polymorphisms include a guanine to adenine change in the 5' upstream regulatory element at residue 84 of SEQ ID NO:1 (GI 187166) and a guanine to adenine change in the 5' upstream regulatory element at residue 137 of SEQ ID NO:1 (GI 187166) (In, *et al.* (1997) *J. Clinical Invest.* 99:1130-1137). These known polymorphisms are also listed in Table 1 under "Known SNPs" and correspond to SEQ ID NOs.:7 and 8. These known polymorphisms may be used in combination with the newly identified polymorphisms of the present invention in the diagnostic, prognostic, pharmacogenomic and therapeutic methods described herein.

The haplotype of the invention is in linkage disequilibrium with known polymorphisms in the promoter region of the 5-LO gene. These polymorphisms are variations at the Sp1 binding site as described in U.S. Patent Number 6,090,547 and In, *et al.* ((1997) *J. Clin. Invest.* 99(5):1130-1137). The wild-type 5-LO gene has a Sp1 binding site comprising five Sp1 motifs in tandem. Each of the Sp1 motifs consists of the nucleotide sequence "GGCGGG" (SEQ ID NO:61). The wild-type Sp1 site is located at nucleotide residues 1670-1699 in the reference sequence GI 187166 (SEQ ID NO:1). Polymorphisms contained within the Sp1 binding site repeat which are referred to herein as "variant Sp1

binding sites" have the deletion of one, the deletion of two, or the addition of one, Sp1 binding motif in the promoter region of the 5-LO gene resulting in three, four, or six Sp1 motifs.

The nucleic acid molecules of the invention can be double- or single-stranded. Accordingly, in one embodiment of the invention, a complement of the nucleotide sequence is provided wherein the polymorphism has been identified. For example, where there has been a single nucleotide change from a guanine to an adenine in a single strand, the complement of that strand will contain a change from cytidine to thymine at the corresponding nucleotide residue.

Nucleic acids of the invention can function as probes or primers, *e.g.,* in methods for determining the allelic identity of a 5-LO polymorphic region in a nucleic acid of interest. The invention further provides vectors comprising the nucleic acid molecules of the present invention; host cells transfected with said vectors whether prokaryotic or eukaryotic; and transgenic non-human animals which contain a heterologous form of a functional or non-functional 5-LO allele described herein. Such a transgenic animal can serve as an animal model for studying the effect of specific 5-LO allelic variations, including mutations, as well as for use in drug screening and/or recombinant protein production.

The invention further provides methods for determining the molecular structure of at least a portion of a 5-LO gene. In a preferred embodiment, the method comprises contacting a sample nucleic acid comprising a 5-LO gene sequence with a probe or primer having a sequence which is complementary to a 5-LO gene sequence, carrying out a reaction that would amplify and/or detect differences in a region of interest within the 5-LO gene sequence, and comparing the result of each reaction with that of a reaction with a control (known) 5-LO gene (*e.g.,* a 5-LO gene from a human not afflicted with an inflammatory condition *e.g.,* asthma, or another disease associated with an aberrant 5-LO activity) so as to determine the molecular structure of the 5-LO gene sequence in the sample nucleic acid. The method of the invention can be used for example in determining the molecular structure of at least a portion of an exon, an intron, a 5' upstream regulatory element, or the 3' untranslated region. In a preferred embodiment, the method comprises determining the identity of at least one nucleotide. In another preferred embodiment, the method comprises determining the

nucleotide at residue 1000 of the reference sequence GI 8247778, any one of residues 472-477 of the reference sequence GI 187166, and/or residue 559 of the reference sequence GI 187166, and combinations thereof. In another embodiment, the method comprises determining the nucleotide at residue 1000 of the reference sequence GI 8247778, any one of residues 472-477 of the reference sequence GI 187166, and/or residue 559 of the reference sequence GI 187166, the nucleotide at residue 84 of GI 187166, and/or the nucleotide at residue 137 of GI 187166. In a further embodiment, the method comprises determining the nucleotide at residue 1000 of the reference sequence GI 8247778, any one of residues 472-477 of the reference sequence GI 187166, and/or residue 559 of the reference sequence GI 187166, and the number of Sp1 motifs in the promoter region of the 5-LO gene.

In another preferred embodiment, the method comprises determining the nucleotide content of at least a portion of a 5-LO gene, such as by sequence analysis. In yet another embodiment, determining the molecular structure of at least a portion of a 5-LO gene is carried out by single-stranded conformation polymorphism (SSCP). In yet another embodiment, the method is an oligonucleotide ligation assay (OLA). Other methods within the scope of the invention for determining the molecular structure of at least a portion of a 5-LO gene include hybridization of allele-specific oligonucleotides, sequence specific amplification, primer specific extension, and denaturing high performance liquid chromatography (DHPLC). In at least some of the methods of the invention, the probe or primer is allele specific. Preferred probes or primers are single stranded nucleic acids, which optionally are labeled.

The invention further provides forensic methods based on detection of polymorphisms in the 5-LO gene.

The invention also provides probes and primers comprising oligonucleotides which hybridizes to at least 6 consecutive nucleotides of any of the sequences set forth in SEQ ID NOs: 4-6, or to the complement of any of such sequences, or naturally occurring mutants or variants thereof. In preferred embodiments, the probe/primer further includes a label attached thereto, which is capable of being detected.

In another embodiment, the invention provides a kit for amplifying and/or for determining the molecular structure of at least a portion of a 5-LO gene, comprising a probe

or primer capable of hybridizing to a 5-LO gene and instructions for use. In one embodiment, the probe or primer is capable of hybridizing to a 5-LO intron. In another embodiment, the probe or primer is capable of hybridizing to a 5-LO allelic variant, preferrably a variant corresponding to 5 loprr1, 5lo01a or 5lo04a. In another preferred embodiment, the polymorphic region is located in the 5' upstream regulatory element. In a preferred embodiment, determining the molecular structure of a region of a 5-LO gene comprises determining the identity of the allelic variant of the polymorphic region. Determining the molecular structure of at least a portion of a 5-LO gene can comprise determining the identity of at least one nucleotide or determining the nucleotide composition, *e.g.*, the nucleotide sequence.

A kit of the invention can be used, *e.g.*, for determining whether a patient will or will not be responsive to effective treatment of a disease associated with a specific 5-LO allelic variant, *e.g.*, asthma, with a 5-LO inhibitor. In a preferred embodiment, the invention provides a kit for determining whether a patient will or will not be responsive to treatment of a inflammatory or allergic disease or condition associated with abnormal leukotriene synthesis, for example, asthma. The kit of the invention can also be used in selecting the appropriate drug to administer to a patient to treat such a disease or condition.

In another aspect, the invention provides a kit for determining whether a inflammatory or allergic disease patient has a more moderate or more severe disease phenotype associated with a specific 5-LO allelic variant of a polymorphic region. In one embodiment, the disease or disorder is characterized by an abnormal 5-LO activity, *e.g.*, aberrant 5-LO expression. In another embodiment, the disease or disorder is characterized by an abnormal eosinophil levels.

Other features and advantages of the invention will be apparent from the following detailed description and claims.

Brief Description of the Figures

Figure 1 depicts the nucleotide sequence corresponding to reference sequence GI 187166 (SEQ ID NO:1).

Figure 2 depicts the nucleotide sequence corresponding to reference sequence GI 8247778 (SEQ ID NO:2).

Figure 3 depicts the nucleotide sequence corresponding to reference sequence GI 4502056 (SEQ ID NO:3).

Figure 4 depicts the nucleotide sequence corresponding to reference sequence GI 9887688 (SEQ ID NO:63).

Detailed Description of the Invention

The present invention is based, in part, on the identification of polymorphisms in the 5-LO gene which are associated with inflammatory or allergic disease phenotypes, including the asthma phenotype. The present invention is based, also in part, on the discovery of a haplotype comprising five polymorphisms within the promoter region of the 5-LO gene. As used herein, the term "haplotype" refers to a set of polymorphisms which are in linkage disequilibrium with each other. That is, the polymorphisms comprising the haplotype segregate together. The haplotype identified herein comprises novel polymorphisms and previously known polymorphisms as described herein and as set forth in Table 1.

The haplotype of the invention also has been found to be in linkage disequilibrium with any one of the three variant Sp1 binding sites disclosed in U.S Patent 6,090,047. Ethnic populations differ in the linked variant Sp1 binding site. In Caucasian and Chinese populations, the haplotype is in linkage disequilibrium with the variant Sp1 binding site containing four Sp1 motifs. Whereas in the African American population, the haplotype is in linkage disequilibrium with any one of the three variant Sp1 binding sites (*i.e.*, variant Sp1 binding sites having 3, 4 or 6 Sp1 motifs).

It has been determined that the presence of a variant Sp1 binding site in the 5-LO promoter region reduced 5-LO gene expression (U.S. Patent Number 6,090,547). Therefore, given that the haplotype of invention is linked to a variant Sp1 binding site, the haplotype and its constituent polymorphisms are all markers of reduced 5-LO gene expression. That is, detection of the haplotype or any of its constituent polymorphims can be used to predict whether or not a patient has decreased 5-LO gene expression as compared to patients homozygous for the wild-type 5-LO gene.

The present invention also relates to the discovery that asthmatic individual heterozygous or homozygous for a 5-LO allele comprising the 5lo01a polymorphism have significantly decreased eosinophil levels as compared to related family members having no such allele. While not intending to be bound by any specific theory, the association between the 5lo01a polymorphism and reduced eosinophil levels may be explained by recent studies that showed that some metabolites of 5-LO can induce production of eosinophils (Uraski, *et al.* (2001) *J. Leukocyte Biol.* 69(1):105-112 and Powell *et al.*, (2001) *J. Allergy. Clin. Immunol.* 107(2):272-278). Specifically, where an allele comprises the 5lo01a polymorphism, the allele also necessarily comprises the haplotype of the inventon and a variant Sp1 binding site. 5-LO alleles comprising a variant Sp1 binding site have reduced levels of 5-LO expression which, in turn, result in reduced levels of 5-LO metabolites, thereby ultimately resulting in a reduction in the number of activated eosinophils. Thus, the presence of the haplotype or one or more polymorphisms comprising the haplotype can be used to predict whether a patient has a greater chance of having abnormally low eosinophil count. Eosinophil counts are used as an index of asthmatic disease, and reflect an asthma-related phenotype (see, for example, Jatakanon, A., *et al.* (2000) *Am. J. Respir. Crit. Care Med.* 161(1):64; Kamfar, *et al.* (1999) *J. Asthma* 36(2):153; and Lonnkivist, K. *et al.* (2001) *J. Allergy Clin. Immunol.* 107(5):812). High eosinophil counts in asthma patients are associated with a more severe asthma phenotype as compared to the asthma phenotype in asthma patients having low eosinophil counts. Accordingly, determination of any polymorphism comprising the haplotype of the invention can be used to predict whether an asthma patient has or will have a moderate disease phenotype. Moreover, determination of any polymorphism comprising the haplotype can be used to differentiate among asthmatic populations to those having a moderate disease phenotype and those having a relatively more severe disease phenotype. A disease phenotype, *e.g.*, the asthma phenotype, may be defined by established clinical parameters (see, for example, Moffitt *et al.* (1994) *Am. Fam. Phys.* 50:1039).

Pharmacogenetic studies have shown that the genetic background of individuals play a role in determining the response of an individual to a specific drug. Thus, determining the allelic variants of 5-LO polymorphic regions of an individual can be useful in predicting how

an individual will respond to a specific drug, *e.g.*, a drug for treating a disease or disorder associated with aberrant 5-LO activity and/or an inflammatory or allergic disease such as asthma. As described above, specific 5-LO polymorphisms comprising the haplotype of the present invention result in increased or decreased production or expression of the 5-LO polypeptide. Accordingly, the action of a drug necessitating interaction with a 5-LO protein will be different in individuals carrying such a 5-LO allele. Furthermore, identification of polymorphisms in the 5-LO gene which indicate responsiveness to 5-LO inhibitor therapy are beneficial in the treatment of inflammatory diseases such as asthma because they can serve as markers by which the most appropriate treatment for inflammatory diseases or disorders can be identified.

Accordingly, preferred embodiments of the present invention include methods of identifying inflammatory or allergy disease patients who are likely to be more or less responsive to treatment with 5-LO inhibitors (*e.g.*, a 5-LO inhibitor as described in U.S. Patent Number 5,703,093, 5,750,565 or 6,025,384, the contents of which are incorporated herein by reference). For example, the presence of a variant Sp1 binding site indicates decreased expression of the 5-LO gene and decreased responsiveness to treatment with 5-LO inhibitors (see U.S. Patent Number 6,090,547 and In, *et al.* (1997), *supra*). Accordingly, the presence of any one of the polymorphisms comprising the haplotype of the present invention indicates the presence of a variant Sp1 binding site and thus the associated decreased expression of the 5-LO gene and decreased responsiveness to treatment with 5-LO inhibitors.

Therefore, the methods of the invention can be used to predict or determine whether a patient will be more or less responsive to treatment with a 5-LO inhibitor or other compounds or agents which interfere with the leukotriene metabolic pathway. These methods are applicable to all inflammatory and allergic diseases in which treatment with 5-LO inhibitors or related agents is appropriate, and in particular, to asthma.

There are multiple alleles of the 5-LO gene. The reference 5-LO gene sequence designated herein is presumed to be the wild-type 5-LO gene sequence and comprises nucleotide sequences that have been deposited in GenBank™ and assigned the Accession Number GI 187166, GI 8247778, and GI 4502056 (corresponding to SEQ ID NO:1, SEQ ID NO:2, and SEQ ID NO:3, respectively). The present invention relates to variant alleles of

the 5-LO gene that differ from the reference 5-LO gene sequence by at least one of the polymorphisms identified in Table 1, and those in linkage equilibrium therewith. The present invention thus relates to nucleic acids comprising such variant 5-LO alleles.

The invention further relates to nucleic acids comprising portions of such variant 5-LO alleles that contain the haplotype of the invention or any of the novel 5-LO polymorphisms identified in Table 1 and are at least 5 nucleotides or basepairs in length. Portions can be, for example, 5-10, 5-15, 10-20, 2-25, 10-30, 10-50 or 10-100 nucleotides or basepairs long. For example, a portion of a variant allele which is 21 nucleotides or basepairs in length includes a 5-LO polymorphism (*i.e.,* a nucleotide which differs from the corresponding nucleotide in the reference 5-LO sequence) and twenty additional nucleotides or basepairs which flank the polymorphism in the variant allele. These additional nucleotides and basepairs can be on one or both sides of the polymorphism. Polymorphisms of the invention are defined in Table 1 with respect to specific reference 5-LO sequences identified in Table 1 (GI 187166, GI 8247778, or GI 4502056). For example, the 5loprr1 polymorphism consists of an adenine substitution for the guanine at residue 1000 of the GI 187166 reference sequence (SEQ ID NO:1).

The 5-LO polymorphisms of the present invention have been identified in the human 5-LO gene by analyzing the DNA of human populations. In particular, DNA samples from 144 individuals were obtained and used for polymorphism discovery. These 144 DNA samples included samples from a population of forty-eight North American Caucasian individuals, forty-eight African American individuals, and forty-eight individuals from throughout the Anhui province in East Central China. The same frequency of 5-LO allelic variants was identified in all three of these groups. Furthermore, the forty-eight DNA samples from the Chinese individuals are from unrelated control patients, enrolled at random. Those patients were assessed for a number of traits related to asthma and allergies. Among the primary variables of interest were lung function, a response to "skin-prick" tests, physician's diagnosis of asthma, total serum IgE, and peripheral blood eosinophils. The peripheral blood eosinophil count was measured using the Coulter counter technique as described in Barnard D.F., *et al.* (1989) *Clin Lab Haematol* 11(3):255-66, and is expressed in terms of eosinophils per microliter.

The allelic variants of the present invention were identified by performing denaturing high performance liquid chromatography (DHPLC) analysis, the polymerase chain reaction (PCR), and/or single stranded conformation polymorphism (SSCP) analysis of genomic DNA from independent individuals as described in Example 1, using PCR primers complementary to intronic sequences surrounding each of the exons, 3' UTR, and 5' upstream regulatory element sequences of the 5-LO gene. The nucleotide sequence of these PCR primers (having SEQ ID NOs.:9-60) is shown in Table 3 (see the Examples).

The presence of three novel polymorphisms in the human 5-LO gene were identified in the populations studied. Two of the polymorphisms were characterized as single nucleotide polymorphisms (SNPs). The remaining polymorphism comprises deletions of two or more nucleotides from the reference 5-LO sequence. This variant is referred to herein as a "deletion variant."

One polymorphism found in the population screened is a change from a guanine to an adenine in the 5' upstream regulatory element of the 5-LO gene at residue 1000 of the reference sequence GI 187166 (polymorphism ID No. 5loprr1). A second polymorphism is a deletion in the 5' upstream regulatory element at residues 472-477 of the reference sequence GI 187166 (polymorphism ID No. 5lo01a). A third polymorphism is a change from a guanine to an adenine in the 5' upstream regulatory element at residue 559 of the reference sequence GI 187166 (polymorphism ID No. 5lo04a). These polymorphisms are listed in Table 1 and correspond to SEQ ID NOs: 4-6.

Additional polymorphisms which are included in the haplotype of the present invention have also been identified in the 5-LO gene. These polymorphisms include a guanine to adenine change in the 5' upstream regulatory element at residue 84 of the reference sequence GI 187166 (polymorphism ID No. 5lonrra) and a guanine to adenine change in the 5' upstream regulatory element at residue 137 of the reference sequence GI 187166 (polymorphism ID No. 5lonrrb (In, *et al.* (1997) *J. Clinical Invest.* 99:1130-1137). These known polymorphisms are also listed in Table 1 under "Known SNPs" and correspond to SEQ ID NOs.:7 and 8. These known polymorphisms may be used in combination with the polymorphisms of the present invention in the diagnostic, prognostic, pharmacogenomic and therapeutic methods described herein.

The haplotype of the present invention were found to be in linkage disequilibrium with polymorphisms that have been identified in the promoter region of the 5-LO gene. These polymorphisms are variations in the Sp1 binding site as described in U.S. Patent Number 6,090,547 and In, *et al.* ((1997) *J. Clin. Invest.* 99(5):1130-1137), the contents of which are incorporated herein by reference. The wild-type 5-LO gene has a Sp1 binding site that comprises five Sp1 motifs in tandem at nucleotide residues 1670-1699 in the reference sequence GI 187166 (SEQ ID NO:1). Each of the motifs has the nucleotide sequence "GGCGGG" (SEQ ID NO:61). Polymorphisms at the Sp1 binding site are referred to herein as "variant Sp1 binding sites" and have the deletion of one or two Sp1 motifs from, or the addition of one Sp1 motif to the wild-type Sp1 binding site, resulting variant Sp1 binding sites containing three, four, or six Sp1 motifs. These known Sp1 binding site polymorphisms may be used in combination with the polymorphisms and haplotype of the present invention in the diagnostic, prognostic, pharmacogenomic and therapeutic methods described herein.

Table 1 contains a "polymorphism ID No." in column 2, which is used herein to identify each individual 5-LO polymorphism. In Table 1, the nucleotide sequence flanking each polymorphism is provided in column 3, wherein the polymorphic residue(s), having the wild-type or reference nucleotide, is indicated in lower-case letters. There are 10 nucleotides flanking the polymorphic nucleotide residue (*i.e.,* 10 nucleotides 5' of the polymorphism and 10 nucleotides 3' of the polymorphism). Column 1 indicates the sequence listing identifier number (SEQ ID NO.) of the sequence shown in column 3 but with a variant nucleotide at the residue(s) shown in lower-case letter(s) or with a deletion of the sequences contained within the parenthesis. For example, SEQ ID NO:4 contains an adenine ("a") at the location indicated by the lower-case letter "g" in the corresponding sequence in column 3. Therefore, SEQ ID NO:4 is identical to the corresponding sequence in column 3, except that the "g" (guanine) residue is replaced by an "a" (adenine residue). The deletion variant, polymorphism ID Nos. 5lo1a, lacks the deleted residues, which are identified by parenthesis in column 3. In Table 1, the location of some primers that may be used to detect each polymorphism is indicated in column 4. For example, primer location "2" indicates that the primers were designed to flank exon 2 and a portion of the surrounding intronic sequence. In column 5, the location of the polymorphism within the 5-LO gene is also provided,

identifying the polymorphism as located in the 5' upstream regulatory element, the 3' untranslated region, or in an intron or an exon. Column 5 also identifies, for each SNP or deletion, the specific nucleotide change (*i.e.*, "5' upstream regulatory element G/A" refers to a single nucleotide polymorphism within the 5' upstream regulatory element resulting in a change from a guanine in the reference sequence to an adenine in the variant allele).

Each polymorphism is identified based on a change in the nucleotide sequence from a "reference sequence." As used herein, the reference sequence is the nucleotide sequence of SEQ ID NO:1 which corresponds to GI 187166 (see Figure 1), the nucleotide sequence of SEQ ID NO:2 which corresponds to GI 247778 (see Figure 2), or the nucleotide sequence of SEQ ID NO:3, which corresponds to GI 4502056 (see Figure 3).

To identify the location of each polymorphism in Table 1, a specific nucleotide residue in a reference sequence is listed for each polymorphism, where nucleotide residue number 1 is the first (*i.e.*, 5') nucleotide in GI 187166 (corresponding to SEQ ID NO:1), the first nucleotide in GI 8247778 (corresponding to SEQ ID NO:2) or the first nucleotide in GI 450205 (corresponding to SEQ ID NO:3). Column 6 lists the reference sequence and polymorphic residue(s) for each polymorphism. For example, polymorphism ID No. 5loprr1 is a single nucleotide polymorphism (SNP) in the 5' upstream regulatory element of the 5-LO gene which results in a change from a guanine to an adenine in the reference sequence GI 187166 (corresponding to SEQ ID NO:1). The SNP is located at residue number 1000 in the reference sequence (SEQ ID NO:1) starting from residue 1 at the 5' end of the nucleotide sequence.

As can be seen in Table 1, one polymorphism found in the population is a change from a guanine to an adenine in the 5' upstream regulatory element region of the 5-LO gene at residue 1000 of GI 18766 (polymorphism ID No. 5loprr1) (SEQ ID NO:4). A second polymorphism is a deletion in the 5' upstream regulatory element at residues 472-477 of GI 18766 (polymorphism ID No. 5lo01a) (SEQ ID NO:5). A third polymorphism is a change from a guanine to an adenine in the 5' upstream regulatory element at residue 559 of GI 18766 (polymorphism ID No. 5lo04a) (SEQ ID NO:6). The nucleic acid molecules of the invention can be double- or single-stranded. Accordingly, the invention further provides for

the nucleic acid strands comprising sequences complementary to the sequences listed in column 3 of Table 1.

The invention further provides allele-specific oligonucleotides that hybridize to a gene comprising a polymorphism of the invention. Such oligonucleotides will hybridize to one polymorphic form of the nucleic acid molecules described herein but not to the other polymorphic form(s) of the sequence. Thus such oligonucleotides can be used to determine the presence or absence of particular alleles of the polymorphic sequences described herein. These oligonucleotides can be probes or primers.

Not only does the present invention provide for polymorphisms in linkage disequilibrium with the polymorphisms of Table 1, it also provides methods for revealing the existence of yet other polymorphisms in the human 5-LO gene. For example, the polymorphism studies described herein can also be applied to populations in which other inflammatory diseases or allergic disease or disorders are prevalent.

Other aspects of the invention are described below or will be apparent to one of skill in the art in light of the present disclosure.

Definitions

For convenience, the meaning of certain terms and phrases employed in the specification, examples, and appended claims are provided below.

The term "inflammatory disease or allergic disease or disorder" as used herein refers to any disease or disorder characterized by an aberrant inflammatory response. An aberrant inflammatory response includes, for example, abnormal (*e.g.*, an increased or decreased) production of leukotriene molecules. Examples of inflammatory or allergic diseases or disorders include, but are not limited to, asthma, bronchitis, sinusitis, ulcerative colitis, nephritis, amyloidosis, rheumatoid arthritis, sarcoidosis, scleroderma, lupus, non-allergic rhinitis, polymyositis, Reiter's syndrome, psoriasis, pelvic inflammatory disease, orbital inflammatory disease, thrombotic disease, and inappropriate allergic responses to environmental stimuli such as poison ivy, pollen, insect stings and certain foods, including atopic dermatitis and contact dermatitis.

The term "allele," which is used interchangeably herein with "allelic variant" and "variant allele", refers to alternative forms of a gene or portions thereof. Alleles occupy the same locus or position on homologous chromosomes. When a patient has two identical alleles of a gene, the patient is said to be homozygous for the gene or allele. When a patient has two different alleles of a gene, the patient is said to be heterozygous for the gene. Alleles of a specific gene, including 5-LO, can differ from each other in a single nucleotide, or several nucleotides, and can include substitutions, deletions, and insertions of nucleotides. An allele of a gene can also be a form of a gene containing one or more mutations.

The term "allelic variant of a 5-LO gene" or "5-LO allelic variant" refers to an alternative form of the 5-LO gene having one of several possible nucleotide sequences found in same position within the gene in the population. The predominate alleles in the population are referred to as "wild-type" alleles.

"Biological activity" or "bioactivity" or "activity" or "biological function", which are used interchangeably, for the purposes herein when applied to 5-LO, means an effector or antigenic function that is directly or indirectly performed by a 5-LO polypeptide (whether in its native or denatured conformation), or by a fragment thereof. Biological activities include modulation of the stereo-specific addition of molecular oxygen to arachidonic acid to form leukotriene A4, thereby modulating synthesis of leukotrienes and modulating inflammatory response, and other biological activities, whether presently known or inherent. A 5-LO bioactivity can be modulated by directly affecting a 5-LO protein effected by, for example, changing the level of effector or substrate level. Alternatively, a 5-LO bioactivity can be modulated by modulating the level of a 5-LO protein, such as by modulating expression of a 5-LO gene. Antigenic functions include possession of an epitope or antigenic site that is capable of cross-reacting with antibodies that bind a native or denatured 5-LO polypeptide or fragment thereof.

Biologically active 5-LO polypeptides include polypeptides having both an effector and antigenic function, or only one of such functions. 5-LO polypeptides include antagonist polypeptides and native 5-LO polypeptides, provided that such antagonists include an epitope of a native 5-LO polypeptide. An effector function of 5-LO polypeptide can be the ability to bind to a ligand, *e.g.*, an arachidonic acid or modified form thereof.

As used herein the term "bioactive fragment of a 5-LO protein" refers to a fragment of a full-length 5-LO protein, wherein the fragment specifically mimics or antagonizes the activity of a wild-type 5-LO protein. The bioactive fragment preferably is a fragment capable of binding to a second molecule, such as a ligand.

The term "an aberrant activity" or "abnormal activity", as applied to an activity of a protein such as 5-LO, refers to an activity which differs from the activity of the wild-type (*i.e.*, normal) protein or which differs from the activity of the protein in a healthy subject, *e.g.*, a subject not afflicted with a disease associated with a 5-LO allelic variant. An activity of a protein can be aberrant because it is stronger than the activity of its wild-type counterpart. Alternatively, an activity of a protein can be aberrant because it is weaker or absent relative to the activity of its wild-type counterpart. An aberrant activity can also be a change in reactivity. For example an aberrant protein can interact with a different protein or ligand relative to its wild-type counterpart. A cell can also have aberrant 5-LO activity due to overexpression or underexpression of the 5-LO gene. Aberrant 5-LO activity can result from a mutation in the gene, which results, *e.g.*, in lower or higher binding affinity of a lipid to the 5-LO protein encoded by the mutated gene. Aberrant 5-LO activity can also result from a lower or higher level of 5-LO receptor on cells, which can result, *e.g.*, from a mutation in the 5' flanking region of the 5-LO gene or any other regulatory element of the 5-LO gene, such as a regulatory element located in an intron. Accordingly, aberrant 5-LO activity can result from an abnormal 5-LO 5' upstream regulatory element activity.

The terms "abnormal 5-LO 5' upstream regulatory element activity", "aberrant 5-LO 5' upstream regulatory element activity", "abnormal 5-LO promoter activity", "aberrant 5-LO promoter activity", "abnormal 5-LO transcriptional activity" and "aberrant 5-LO transcriptional activity", which are used interchangeably herein, refer to the transcriptional activity of a 5-LO 5' upstream regulatory element which differs from the transcriptional activity of the corresponding 5' upstream regulatory element in the wild-type 5-LO allele. Abnormal 5-LO activity can result from a higher or lower transcriptional activity as compared to transcriptional activity of a wild-type 5-LO allele. Aberrant 5-LO 5' upstream regulatory element activity can result, for example, from the presence of a genetic lesion in a regulatory element, such as in a 5' upstream regulatory element. An "aberrant 5-LO 5'

upstream regulatory element activity" is also intended to refer to the transcriptional activity of a 5-LO 5' upstream regulatory element which is functional (capable of inducing transcription of a gene to which it is operably linked) in tissues or cells in which the normal or wild-type 5-LO 5' upstream regulatory element is not functional or which is non functional in tissues or cells in which the normal or wild-type 5-LO 5' upstream regulatory element is functional. Thus, a tissue distribution of 5-LO in a patient which differs from the tissue distribution of 5-LO in a normal (*e.g.*, healthy) individual, can be the result of abnormal transcriptional activity from the 5-LO 5' upstream regulatory element. Such abnormal transcriptional activity can result from *e.g.*, one or more mutations in a regulatory element, such as in a 5' upstream regulatory element thereof. Abnormal transcriptional activity can also result from a mutation in a transcription factor involved in the control of 5-LO gene expression.

"Cells," "host cells" or "recombinant host cells" are terms used interchangeably herein. It is understood that such terms refer not only to the particular patient cell but to the progeny or derivatives of such a cell. Because certain modifications may occur in succeeding generations due to either mutation or environmental influences, such progeny may not, in fact, be identical to the parent cell, but are still included within the scope of the term as used herein.

As used herein, the term "gene" or "recombinant gene" refers to a nucleic acid molecule comprising an open reading frame and including at least one exon and (optionally) an intron sequence. The term "intron" refers to a DNA sequence present in a given gene which is spliced out during mRNA maturation.

"Homology" or "identity" or "similarity" refers to sequence similarity between two peptides or between two nucleic acid molecules. Homology can be determined by comparing a position in each sequence which may be aligned for purposes of comparison. When a position in the compared sequence is occupied by the same base or amino acid, then the molecules are homologous at that position. A degree of homology between sequences is a function of the number of matching or homologous positions shared by the sequences. An "unrelated" or "non-homologous" sequence shares less than 40 % identity, though preferably less than 25 % identity, with one of the sequences of the present invention.

To determine the percent identity of two amino acid sequences or of two nucleic acids, the sequences are aligned for optimal comparison purposes (*e.g.*, gaps can be introduced in the sequence of a first amino acid or nucleic acid sequence for optimal alignment with a second amino or nucleic acid sequence). The amino acid residues or nucleotides at corresponding amino acid positions or nucleotide positions are then compared. When a position in the first sequence is occupied by the same amino acid residue or nucleotide as the corresponding position in the second sequence, then the molecules are identical at that position. The percent identity between the two sequences is a function of the number of identical positions shared by the sequences (*i.e.*, % identity = number of identical positions/total number of positions (*e.g.*, overlapping positions) x100). In one embodiment the two sequences are the same length.

The determination of percent identity between two sequences can be accomplished using a mathematical algorithm. A preferred, non-limiting example of a mathematical algorithm utilized for the comparison of two sequences is the algorithm of Karlin and Altschul (1990) *Proc. Natl. Acad. Sci. USA* 87:2264-2268, modified as in Karlin and Altschul (1993) *Proc. Natl. Acad. Sci. USA* 90:5873-5877. Such an algorithm is incorporated into the NBLAST and XBLAST programs of Altschul, *et al.* (1990) *J. Mol. Biol.* 215:403-410. BLAST nucleotide searches can be performed with the NBLAST program, score = 100, wordlength = 12 to obtain nucleotide sequences homologous to a nucleic acid molecules of the invention. BLAST protein searches can be performed with the XBLAST program, score = 50, wordlength = 3 to obtain amino acid sequences homologous to a protein molecules of the invention. To obtain gapped alignments for comparison purposes, Gapped BLAST can be utilized as described in Altschul *et al.* (1997) *Nucleic Acids Res.* 25:3389-3402. Alternatively, PSI-Blast can be used to perform an iterated search which detects distant relationships between molecules. When utilizing BLAST, Gapped BLAST, and PSI-Blast programs, the default parameters of the respective programs (*e.g.*, XBLAST and NBLAST) can be used. Another preferred, non-limiting example of a mathematical algorithm utilized for the comparison of sequences is the algorithm of Myers and Miller, (1988) *CABIOS* 4:11-17. Such an algorithm is incorporated into the ALIGN program (version 2.0) which is part of the GCG sequence alignment software package. When

utilizing the ALIGN program for comparing amino acid sequences, a PAM120 weight residue table, a gap length penalty of 12, and a gap penalty of 4 can be used. Yet another useful algorithm for identifying regions of local sequence similarity and alignment is the FASTA algorithm as described in Pearson and Lipman (1988) *Proc. Natl. Acad. Sci. USA* 85:2444-2448. When using the FASTA algorithm for comparing nucleotide or amino acid sequences, a PAM120 weight residue table can, for example, be used with a *k*-tuple value of 2.

The term "a homolog of a nucleic acid" refers to a nucleic acid having a nucleotide sequence having a certain degree of homology with the nucleotide sequence of the nucleic acid or complement thereof. For example, a homolog of a double stranded nucleic acid having SEQ ID NO:N is intended to include nucleic acids having a nucleotide sequence which has a certain degree of homology with SEQ ID NO:N or with the complement thereof. Preferred homologs of nucleic acids are capable of hybridizing to the nucleic acid or complement thereof.

The term "hybridization probe" or "primer" as used herein is intended to include oligonucleotides which hybridize bind in a base-specific manner to a complementary strand of a target nucleic acid. Such probes include peptide nucleic acids, and described in Nielsen *et al.,* (1991) *Science* 254:1497-1500. Probes and primers can be any length suitable for specific hybridization to the target nucleic acid sequence. The most appropriate length of the probe and primer may vary depending on the hybridization method in which it is being used; for example, particular lengths may be more appropriate for use in microfabricated arrays, while other lengths may be more suitable for use in classical hybridization methods. Such optimizations are known to the skilled artisan. Suitable probes and primers can range form about 5 nucleotides to about 30 nucleotides in length. For example, probes and primers can be 5, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 25, 26, 28 or 30 nucleotides in length. The probe or primer of the invention comprises a sequence that flanks and/or preferably overlaps, at least one polymorphic site occupied by any of the possible variant nucleotides. The nucleotide sequence of an overlapping probe or primer can correspond to the coding sequence of the allele or to the complement of the coding sequence of the allele.

The term "interact" as used herein is meant to include detectable interactions between molecules, such as can be detected using, for example, a binding or hybridization assay. The term interact is also meant to include "binding" interactions between molecules. Interactions may be, for example, protein-protein, protein-nucleic acid, protein-small molecule or small molecule-nucleic acid in nature.

The term "intronic sequence" or "intronic nucleotide sequence" refers to the nucleotide sequence of an intron or portion thereof.

The term "isolated" as used herein with respect to nucleic acids, such as DNA or RNA, refers to molecules separated from other DNAs or RNAs, respectively, that are present in the natural source of the macromolecule. The term isolated as used herein also refers to a nucleic acid or peptide that is substantially free of cellular material, viral material, or culture medium when produced by recombinant DNA techniques, or chemical precursors or other chemicals when chemically synthesized. Moreover, an "isolated nucleic acid" is meant to include nucleic acid fragments which are not naturally occurring as fragments and would not be found in the natural state. The term "isolated" is also used herein to refer to polypeptides which are isolated from other cellular proteins and is meant to encompass both purified and recombinant polypeptides.

The term "linkage" describes the tendency of genes, alleles, loci or genetic markers to be inherited together as a result of their location on the same chromosome. It can be measured by percent recombination between the two genes, alleles, loci, or genetic markers. The term "linkage disequilibrium" refers to a greater than random association between specific alleles at two marker loci within a particular population. In general, linkage disequilibrium decreases with an increase in physical distance. If linkage disequilibrium exists between two markers within one gene, then the genotypic information at one marker can be used to make probabilistic predictions about the genotype of the second marker.

The term "locus" refers to a specific position in a chromosome. For example, a locus of a 5-LO gene refers to the chromosomal position of the 5-LO gene.

The term "modulation" as used herein refers to both upregulation, (*i.e.*, activation or stimulation), for example by agonizing; and downregulation (*i.e.* inhibition or suppression), for example by antagonizing of a bioactivity (*e.g.* expression of a gene).

The term "molecular structure" of a gene or a portion thereof refers to the structure as defined by the nucleotide content (including deletions, substitutions, additions of one or more nucleotides), the nucleotide sequence, the state of methylation, and/or any other modification of the gene or portion thereof.

The term "mutated gene" refers to an allelic form of a gene that differs from the predominant form in a population. A mutated gene is capable of altering the phenotype of a patient having the mutated gene relative to a patient having the predominant form of the gene. If a patient must be homozygous for this mutation to have an altered phenotype, the mutation is said to be recessive. If one copy of the mutated gene is sufficient to alter the phenotype of the patient, the mutation is said to be dominant. If a patient has one copy of the mutated gene and has a phenotype that is intermediate between that of a homozygous and that of a heterozygous (for that gene) patient, the mutation is said to be co-dominant.

As used herein, the term "nucleic acid" refers to polynucleotides such as deoxyribonucleic acid (DNA), and, where appropriate, ribonucleic acid (RNA). The term should also be understood to include, as equivalents, derivatives, variants and analogs of either RNA or DNA made from nucleotide analogs, and, as applicable to the embodiment being described, single (sense or antisense) and double-stranded polynucleotides. Deoxyribonucleotides include deoxyadenosine, deoxycytidine, deoxyguanosine, and deoxythymidine. For purposes of clarity, when referring herein to a nucleotide of a nucleic acid, which can be DNA or an RNA, the terms "adenosine", "cytidine", "guanosine", and thymidine" are used. It is understood that if the nucleic acid is RNA, a nucleotide having a uracil base is uridine.

The term "nucleotide sequence complementary to the nucleotide sequence set forth in SEQ ID NO:N" refers to the nucleotide sequence of the complementary strand of a nucleic acid strand having SEQ ID NO:N. The term "complementary strand" is used herein interchangeably with the term "complement". The complement of a nucleic acid strand can be the complement of a coding strand or the complement of a non-coding strand. When referring to double stranded nucleic acids, the complement of a nucleic acid having SEQ ID NO:N refers to the complementary strand of the strand having SEQ ID NO:N or to any nucleic acid having the nucleotide sequence of the complementary strand of SEQ ID NO:N.

When referring to a single stranded nucleic acid having the nucleotide sequence SEQ ID NO:N, the complement of this nucleic acid is a nucleic acid having a nucleotide sequence which is complementary to that of SEQ ID NO:N. The nucleotide sequences and complementary sequences thereof are always given in the 5' to 3' direction. The term "complement" and "reverse complement" are used interchangeably herein.

A "non-human animal" of the invention can include mammals such as rodents, non-human primates, sheep, goats, horses, dogs, cows, chickens, amphibians, reptiles, etc. Preferred non-human animals are selected from the rodent family including rat and mouse, most preferably mouse, though transgenic amphibians, such as members of the *Xenopus* genus, and transgenic chickens can also provide important tools for understanding and identifying agents which can affect, for example, embryogenesis and tissue formation. The term "chimeric animal" is used herein to refer to animals in which an exogenous sequence is found, or in which an exogenous sequence is expressed in some but not all cells of the animal. The term "tissue-specific chimeric animal" indicates that an exogenous sequence is present and/or expressed or disrupted in some tissues, but not others.

The term "oligonucleotide" is intended to include and single- or double stranded DNA or RNA. Oligonucleotides can be naturally occurring or synthetic, but are typically prepared by synthetic means. Preferred oligonucleotides of the invention include segments of 5-LO gene sequence or their complements, which include and/or flank any one of the polymorphic sites shown in Table 1. The segments can be between 5 and 250 bases, and, in specific embodiments, are between 5-10, 5-20, 10-20, 10-50, 20-50 or 10-100 bases. For example, the segments can be 21 bases. The polymorphic site can occur within any position of the segment or a region next to the segment. The segments can be from any of the allelic forms of 5-LO gene sequence shown in Table 1.

The term "operably-linked" is intended to mean that the 5' upstream regulatory element is associated with a nucleic acid in such a manner as to facilitate transcription of the nucleic acid from the 5' upstream regulatory element.

The term "polymorphism" refers to the coexistence of more than one form of a gene or portion thereof. A portion of a gene of which there are at least two different forms, *i.e.*, two different nucleotide sequences, is referred to as a "polymorphic region of a gene". A

polymorphic locus can be a single nucleotide, the identity of which differs in the other alleles. A polymorphic locus can also be more than one nucleotide long. The allelic form occurring most frequently in a selected population is often referred to as the reference and/or wild-type form. Other allelic forms are typically designated or alternative or variant alleles. Diploid organisms may be homozygous or heterozygous for allelic forms. A diallelic or biallelic polymorphism has two forms. A trialleleic polymorphism has three forms.

A "polymorphic gene" refers to a gene having at least one polymorphic region.

The term "primer" as used herein, refers to a single-stranded oligonucleotide which acts as a point of initiation of template-directed DNA synthesis under appropriate conditions (*e.g.*, in the presence of four different nucleoside triphosphates and as agent for polymerization, such as DNA or RNA polymerase or reverse transcriptase) in an appropriate buffer and at a suitable temperature. The length of a primer may vary but typically ranges from 15 to 30 nucleotides. A primer need not match the exact sequence of a template, but must be sufficiently complementary to hybridize with the template.

The term "primer pair" refers to a set of primers including an upstream primer that hybridizes with the 3' end of the complement of the DNA sequence to be amplified and a downstream primer that hybridizes with the 3' end of the sequence to be amplified.

The terms "protein", "polypeptide" and "peptide" are used interchangeably herein when referring to a gene product.

The term "recombinant protein" refers to a polypeptide which is produced by recombinant DNA techniques, wherein generally, DNA encoding the polypeptide is inserted into a suitable expression vector which is in turn used to transform a host cell to produce the heterologous protein.

A "regulatory element", also termed herein "regulatory sequence" is intended to include elements which are capable of modulating transcription from a 5' upstream regulatory sequence, including, but not limited to a basic promoter, and include elements such as enhancers and silencers. The term "enhancer", also referred to herein as "enhancer element", is intended to include regulatory elements capable of increasing, stimulating, or enhancing transcription from a 5' upstream regulatory element, including a basic promoter. The term "silencer", also referred to herein as "silencer element" is intended to include

regulatory elements capable of decreasing, inhibiting, or repressing transcription from a 5' upstream regulatory element, including a basic promoter. Regulatory elements are typically present in 5' flanking regions of genes. Regulatory elements also may be present in other regions of a gene, such as introns. Thus, it is possible that 5-LO genes have regulatory elements located in introns, exons, coding regions, and 3' flanking sequences. Such regulatory elements are also intended to be encompassed by the present invention and can be identified by any of the assays that can be used to identify regulatory elements in 5' flanking regions of genes.

The term "regulatory element" further encompasses "tissue specific" regulatory elements, *i.e.,* regulatory elements which effect expression of an operably linked DNA sequence preferentially in specific cells (*e.g.,* cells of a specific tissue). Gene expression occurs preferentially in a specific cell if expression in this cell type is significantly higher than expression in other cell types. The term "regulatory element" also encompasses non-tissue specific regulatory elements, *i.e.,* regulatory elements which are active in most cell types. Furthermore, a regulatory element can be a constitutive regulatory element, *i.e.,* a regulatory element which constitutively regulates transcription, as opposed to a regulatory element which is inducible, *i.e.,* a regulatory element which is active primarily in response to a stimulus. A stimulus can be, *e.g.,* a molecule, such as a protein, hormone, cytokine, heavy metal, phorbol ester, cyclic AMP (cAMP), or retinoic acid.

Regulatory elements are typically bound by proteins, *e.g.,* transcription factors. The term "transcription factor" is intended to include proteins or modified forms thereof, which interact preferentially with specific nucleic acid sequences, *i.e.,* regulatory elements, and which in appropriate conditions stimulate or repress transcription. Some transcription factors are active when they are in the form of a monomer. Alternatively, other transcription factors are active in the form of a dimer consisting of two identical proteins or different proteins (heterodimer). Modified forms of transcription factors are intended to refer to transcription factors having a postranslational modification, such as the attachment of a phosphate group. The activity of a transcription factor is frequently modulated by a postranslational modification. For example, certain transcription factors are active only if they are phosphorylated on specific residues. Alternatively, transcription factors can be active in the

absence of phosphorylated residues and become inactivated by phosphorylation. A list of known transcription factors and their DNA binding site can be found, *e.g.,* in public databases, *e.g.,* TFMATRIX Transcription Factor Binding Site Profile database.

The term "single nucleotide polymorphism" (SNP) refers to a polymorphic site occupied by a single nucleotide, which is the site of variation between allelic sequences. The site is usually preceded by and followed by highly conserved sequences of the allele (*e.g.*, sequences that vary in less than 1/100 or 1/1000 members of a population). SNP usually arises due to substitution of one nucleotide for another at the polymorphic site. SNPs can also arise from a deletion of a nucleotide or an insertion of a nucleotide relative to a reference allele. Typically the polymorphic site is occupied by a base other than the reference base. For example, where the reference allele contains the base "T" (thymidine) at the polymorphic site, the altered allele can contain a "C" (cytidine), "G" (guanine), or "A" (adenine) at the polymorphic site.

SNP's may occur in protein-coding nucleic acid sequences, in which case they may give rise to a defective or otherwise variant protein, or genetic disease. Such a SNP may alter the coding sequence of the gene and therefore specify another amino acid (a "missense" SNP) or a SNP may introduce a stop codon (a "nonsense" SNP). When a SNP does not alter the amino acid sequence of a protein, the SNP is called "silent." SNP's may also occur in noncoding regions of the nucleotide sequence. This may result in defective protein expression, *e.g.,* as a result of alternative spicing, or it may have no effect.

As used herein, the term "specifically hybridizes" or "specifically detects" refers to the ability of a nucleic acid molecule of the invention to hybridize to at least approximately 6, 12, 20, 30, 40, 50, 60, 70, 80, 90, 100, 110, 120, 130 or 140 consecutive nucleotides of either strand of a 5-LO gene.

"5-LO" refers to a 5-lipoxygenase molecule that catalyzes the stereo-specific addition of molecular oxygen to arachidonic acid to form leukotriene A4 (LTA4) (Silverman, *et al.* (1998) *Clin Exp* Allergy 5:164; In , *et al.* (1997) *J. Clin. Invest.* 99:1130-1137), and therefore functions as a modulator of leukotriene synthesis. It has been found that 5-LO expression is dependent on the presence of a 5' GC-rich transcription factor binding region; this region

contains sequences characteristic of binding motifs, *e.g.,* Sp1 binding motifs (In, *et al., supra).*

The term "5-LO therapeutic" refers to various forms of 5-LO protein or polypeptides, as well as peptidomimetics, nucleic acids, or small molecules, which can modulate at least one activity of a 5-LO protein by mimicking or potentiating (agonizing) or inhibiting (antagonizing) the effects of a naturally-occurring 5-LO protein. A 5-LO therapeutic which mimics or potentiates the activity of a wild-type 5-LO protein is a "5-LO agonist". Conversely, a 5-LO therapeutic which inhibits the activity of a wild-type 5-LO protein is a "5-LO antagonist". 5-LO therapeutics can be used to treat diseases which are associated with a specific 5-LO allele which encodes a protein having an amino acid sequence that differs from that of the wild-type 5-LO protein.

As used herein, the term "transfection" means the introduction of a nucleic acid, *e.g.,* an expression vector, into a recipient cell by nucleic acid-mediated gene transfer. The term "transduction" is generally used herein when the transfection with a nucleic acid is by viral delivery of the nucleic acid. "Transformation", as used herein, refers to a process in which a cell's genotype is changed as a result of the cellular uptake of exogenous DNA or RNA, and, for example, the transformed cell expresses a recombinant form of a polypeptide or, in the case of anti-sense expression from the transferred gene, the expression of a naturally-occurring form of the recombinant protein is disrupted.

As used herein, the term "transgene" refers to a nucleic acid sequence which has been genetic-engineered into a cell. Daughter cells deriving from a cell in which a transgene has been introduced are also said to contain the transgene (unless it has been deleted). A transgene can encode, *e.g.,* a polypeptide, or an antisense transcript, partly or entirely heterologous, *i.e.,* foreign, to the transgenic animal or cell into which it is introduced, or, is homologous to an endogenous gene of the transgenic animal or cell into which it is introduced, but which is designed to be inserted, or is inserted, into the animal's genome in such a way as to alter the genome of the cell into which it is inserted (*e.g.,* it is inserted at a location which differs from that of the natural gene or its insertion results in a knockout). Alternatively, a transgene can also be present in an episome. A transgene can include one or

more transcriptional regulatory sequence and any other nucleic acid, (*e.g.* intron), that may be necessary for optimal expression of a selected nucleic acid.

A "transgenic animal" refers to any animal, preferably a non-human animal, *e.g.* a mammal, bird or an amphibian, in which one or more of the cells of the animal contain heterologous nucleic acid introduced by genetic engineering, such as by transgenic techniques well known in the art. The nucleic acid is introduced into the cell, directly or indirectly by introduction into a precursor of the cell, by way of deliberate genetic manipulation, such as by microinjection or by infection with a recombinant virus. The term genetic manipulation does not include classical cross-breeding, or *in vitro* fertilization, but rather is directed to the introduction of a recombinant DNA molecule. This molecule may be integrated within a chromosome, or it may be extrachromosomally replicating DNA. In the typical transgenic animals described herein, the transgene causes cells to express a recombinant form of one of a protein, *e.g.* either agonistic or antagonistic forms. However, transgenic animals in which the recombinant gene is silent are also contemplated, as for example, the FLP or CRE recombinase dependent constructs described below. Moreover, "transgenic animal" also includes those recombinant animals in which gene disruption of one or more genes is caused by human intervention, including both recombination and antisense techniques.

The term "treatment," or "treating" as used herein, is defined as the application or administration of a therapeutic agent to a subject, implementation of lifestyle changes (*e.g.*, changes in diet or environment), administration of medication, or application or administration of a therapeutic agent to a patient who has a disease or disorder, a symptom of disease or disorder or a predisposition toward a disease or disorder, with the purpose to cure, heal, alleviate, relieve, alter, remedy, ameliorate, improve or affect the disease or disorder, the symptoms of the disease or disorder, or the predisposition toward disease.

As used herein, the term "vector" refers to a nucleic acid molecule capable of transporting or replicating another nucleic acid to which it has been linked. One type of preferred vector is an episome, *i.e.*, a nucleic acid capable of extra-chromosomal replication. Preferred vectors are those capable of autonomous replication and/or expression of nucleic acids to which they are linked. Vectors capable of directing the expression of genes to which

they are operatively-linked are referred to herein as "expression vectors". In general, expression vectors of utility in recombinant DNA techniques are often in the form of "plasmids" which refer generally to circular double stranded DNA circles which, in their vector form are not physically linked to the host chromosome. In the present specification, "plasmid" and "vector" are used interchangeably as the plasmid is the most commonly used form of vector. However, the invention is intended to include such other forms of expression vectors which serve equivalent functions and which become known in the art subsequently hereto.

The term "wild-type allele" refers to an allele of a gene which, when present in two copies in a patient results in a wild-type phenotype. There can be several different wild-type alleles of a specific gene, since certain nucleotide changes in a gene may not affect the phenotype of a patient having two copies of the gene with the nucleotide changes. The terms "wild-type" and "reference sequence" are used interchangeably herein.

Polymorphisms of the Invention

The nucleic acid molecules of the invention include specific 5-LO allelic variants, which differ from the reference sequences set forth in SEQ ID NO:1, SEQ ID NO:2, or SEQ ID NO:3, or at least a portion thereof, having a polymorphic region. The preferred nucleic acid molecules of the present invention comprise 5-LO sequences having one or more of the polymorphisms shown in Table 1, those in linkage disequilibrium therewith, and those included in the haplotype, as described herein. Nucleic acid molecules of the invention can function as probes or primers, *e.g.*, in methods for determining the allelic identity of a 5-LO polymorphic region. The nucleic acids of the invention can also be used to determine whether a patient is or is not at risk of developing a disease associated with a specific allelic variant of a 5-LO polymorphic region, *e.g.*, a disease or disorder associated with an aberrant 5-LO activity. The nucleic acids of the invention can further be used to prepare or express 5-LO polypeptides encoded by specific alleles, such as mutant alleles. Such nucleic acids can be used in gene therapy. Polypeptides encoded by specific 5-LO alleles, such as mutant 5-LO polypeptides, can also be used in therapy or for preparing reagents, *e.g.*, antibodies, for

detecting 5-LO proteins encoded by these alleles. Accordingly, such reagents can be used to detect mutant 5-LO proteins.

As described herein, several allelic variants of human 5-LO genes have been identified. The invention is intended to encompass all of these allelic variants as well as, those in linkage disequilibrium which can be identified, *e.g.*, according to the methods described herein. "Linkage disequilibrium" refers to an association between specific alleles at two marker loci within a particular population. In general, linkage disequilbrium decreases with an increase in physical distance. If linkage disequilbrium exists between two markers within one gene, then the genotypic information at one marker can be used to make predictions about the genotype of the second marker.

The invention also provides isolated nucleic acids comprising at least one polymorphic region of a 5-LO gene having a nucleotide sequence which differs from the reference nucleotide sequence set forth in SEQ ID NO:1, SEQ ID NO:2, or SEQ ID NO:3. Preferred nucleic acids have a variant allele located in the 5' upstream regulatory element of the 5-LO gene. Accordingly, preferred nucleic acids of the invention comprise an adenine at residue 1000 of the 5' upstream regulatory element of GI 187166 (as set forth in SEQ ID NO:4) or a thymine at the corresponding position in an otherwise complementary sequence; a deletion of the nucleotides GTTAAA (SEQ ID NO:62) at residues 472-477 of GI 187166 (as set forth in SEQ ID NO:5) or deletion of the nucleotides TTTAAC at the corresponding positions in an otherwise complementary sequence; and/or an adenine at residue 559 of the 5' upstream regulatory element of GI 187166 (as set forth in SEQ ID NO:6) or a thymine at the corresponding position in an otherwise complementary sequence.

The nucleic acid molecules of the invention can be single stranded DNA (*e.g.*, an oligonucleotide), double stranded DNA (*e.g.*, double stranded oligonucleotide) or RNA. Preferred nucleic acid molecules of the invention can be used as probes or primers. Primers of the invention refer to nucleic acids which hybridize to a nucleic acid sequence which is adjacent to the region of interest or which covers the region of interest and is extended. As used herein, the term "hybridizes" is intended to describe conditions for hybridization and washing under which nucleotide sequences that are significantly identical or homologous to each other remain hybridized to each other. Preferably, the conditions are such that

sequences at least about 70%, more preferably at least about 80%, even more preferably at least about 85% or 90% identical to each other remain hybridized to each other. Such stringent conditions vary according to the length of the involved nucleotide sequence but are known to those skilled in the art and can be found or determined based on teachings in *Current Protocols in Molecular Biology*, Ausubel *et al.*, eds., John Wiley & Sons, Inc. (1995), sections 2, 4 and 6. Additional stringent conditions and formulas for determining such conditions can be found in *Molecular Cloning: A Laboratory Manual*, Sambrook *et al.*, Cold Spring Harbor Press, Cold Spring Harbor, NY (1989), chapters 7, 9 and 11. A preferred, non-limiting example of stringent hybridization conditions for hybrids that are at least basepairs in length includes hybridization in 4X sodium chloride/sodium citrate (SSC), at about 65-70°C (or hybridization in 4X SSC plus 50% formamide at about 42-50°C) followed by one or more washes in 1X SSC, at about 65-70°C. A preferred, non-limiting example of highly stringent hybridization conditions for such hybrids includes hybridization in 1X SSC, at about 65-70°C (or hybridization in 1X SSC plus 50% formamide at about 42-50°C) followed by one or more washes in 0.3X SSC, at about 65-70°C. A preferred, non-limiting example of reduced stringency hybridization conditions for such hybrids includes hybridization in 4X SSC, at about 50-60°C (or alternatively hybridization in 6X SSC plus 50% formamide at about 40-45°C) followed by one or more washes in 2X SSC, at about 50-60°C. Ranges intermediate to the above-recited values, *e.g.*, at 65-70°C or at 42-50°C are also intended to be encompassed by the present invention. SSPE (1xSSPE is 0.15M NaCl, 10mM NaH_2PO_4, and 1.25mM EDTA, pH 7.4) can be substituted for SSC (1xSSC is 0.15M NaCl and 15mM sodium citrate) in the hybridization and wash buffers; washes are performed for 15 minutes each after hybridization is complete.

The hybridization temperature for hybrids anticipated to be less than 50 base pairs in length should be 5-10°C less than the melting temperature (T_m) of the hybrid, where T_m is determined according to the following equations. For hybrids less than 18 base pairs in length, $T_m(°C) = 2(\#\text{ of A + T bases}) + 4(\#\text{ of G + C bases})$. For hybrids between 18 and 49 base pairs in length, $T_m(°C) = 81.5 + 16.6(\log_{10}[Na^+]) + 0.41(\%G+C) - (600/N)$, where N is the number of bases in the hybrid, and $[Na^+]$ is the concentration of sodium ions in the hybridization buffer ($[Na^+]$ for 1xSSC = 0.165 M). It will also be recognized by the skilled

practitioner that additional reagents may be added to hybridization and/or wash buffers to decrease non-specific hybridization of nucleic acid molecules to membranes, for example, nitrocellulose or nylon membranes, including but not limited to blocking agents (*e.g.*, BSA or salmon or herring sperm carrier DNA), detergents (*e.g.*, SDS), chelating agents (*e.g.*, EDTA), Ficoll, PVP and the like. When using nylon membranes, in particular, an additional preferred, non-limiting example of stringent hybridization conditions is hybridization in 0.25-0.5M NaH_2PO_4, 7% SDS at about 65°C, followed by one or more washes at 0.02M NaH_2PO_4, 1% SDS at 65°C, see *e.g.*, Church and Gilbert (1984) *Proc. Natl. Acad. Sci. USA* 81:1991-1995, (or alternatively 0.2X SSC, 1% SDS).

A primer or probe can be used alone in a detection method, or a primer can be used together with at least one other primer or probe in a detection method. Primers can also be used to amplify at least a portion of a nucleic acid. Probes of the invention refer to nucleic acids which hybridize to the region of interest and which are not further extended. For example, a probe is a nucleic acid which specifically hybridizes to a polymorphic region of a 5-LO gene, and which by hybridization or absence of hybridization to the DNA of a patient or the type of hybrid formed will be indicative of the identity of the allelic variant of the polymorphic region of the 5-LO gene.

Numerous procedures for determining the nucleotide sequence of a nucleic acid molecule, or for determining the presence of mutations in nucleic acid molecules include a nucleic acid amplification step, which can be carried out by, *e.g.*, polymerase chain reaction (PCR). Accordingly, in one embodiment, the invention provides primers for amplifying portions of a 5-LO gene, such as portions of exons and/or portions of introns. In a preferred embodiment, the exons and/or sequences adjacent to the exons of the human 5-LO gene will be amplified to, *e.g.*, detect which allelic variant, if any, of a polymorphic region is present in the 5-LO gene of a patient. Preferred primers comprise a nucleotide sequence complementary a specific allelic variant of a 5-LO polymorphic region and of sufficient length to selectively hybridize with a 5-LO gene. In a preferred embodiment, the primer, *e.g.*, a substantially purified oligonucleotide, comprises a region having a nucleotide sequence which hybridizes under stringent conditions to about 6, 8, 10, or 12, preferably 25, 30, 40, 50, or 75 consecutive nucleotides of a 5-LO gene. In an even more preferred

embodiment, the primer is capable of hybridizing to a 5-LO nucleotide sequence and comprises a nucleotide sequence of any sequence set forth in any of SEQ ID NOs:4-6, or complements thereof. For example, primers comprising a nucleotide sequence of at least about 15 consecutive nucleotides, at least about 25 nucleotides or having from about 15 to about 20 nucleotides set forth in any of SEQ ID NOs:4-6 or complement thereof are provided by the invention. Primers having a sequence of more than about 25 nucleotides are also within the scope of the invention. Preferred primers of the invention are primers that can be used in PCR for amplifying each of the exons of a 5-LO gene.

Primers can be complementary to nucleotide sequences located close to each other or further apart, depending on the use of the amplified DNA. For example, primers can be chosen such that they amplify DNA fragments of at least about 10 nucleotides or as much as several kilobases. Preferably, the primers of the invention will hybridize selectively to 5-LO nucleotide sequences located about 150 to about 350 nucleotides apart.

For amplifying at least a portion of a nucleic acid, a forward primer (*i.e.,* 5' primer) and a reverse primer (*i.e.,* 3' primer) will preferably be used. Forward and reverse primers hybridize to complementary strands of a double stranded nucleic acid, such that upon extension from each primer, a double stranded nucleic acid is amplified. A forward primer can be a primer having a nucleotide sequence or a portion of the nucleotide sequence shown in Table 1 (SEQ ID NOs:4-6) or in SEQ ID NOs.: 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45, 47, 49, 51, 53, 55, 57, or 59. A reverse primer can be a primer having a nucleotide sequence or a portion of the nucleotide sequence that is complementary to a nucleotide sequence shown in Table 1 (SEQ ID NOs:4-6) or in SEQ ID NOs.: 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, or 60. Preferred pairs of primers for amplifying each of the exons of human 5-LO are set forth in Table 3 (see Example 2).

Yet other preferred primers of the invention are nucleic acids which are capable of selectively hybridizing to an allelic variant of a polymorphic region of a 5-LO gene. Thus, such primers can be specific for a 5-LO gene sequence, so long as they have a nucleotide sequence which is capable of hybridizing to a 5-LO gene. Preferred primers are capable of specifically hybridizing to any of the allelic variants listed in Table 1 (*i.e.,* sequences

comprising any of SEQ ID NOs:4-6 or a complement thereof). Such primers can be used, *e.g.,* in sequence specific oligonucleotide priming as described further herein.

The 5-LO nucleic acids of the invention can also be used as probes, *e.g.,* in therapeutic and diagnostic assays. For instance, the present invention provides a probe comprising a substantially purified oligonucleotide, which oligonucleotide comprises a region having a nucleotide sequence that is capable of hybridizing specifically to a region of a 5-LO gene which is polymorphic (i.e., sequences comprising any of SEQ ID NOs:4-6 or a complement thereof). In an even more preferred embodiment of the invention, the probes are capable of hybridizing specifically to one allelic variant of a 5-LO gene having a nucleotide sequence which differs from the nucleotide sequence set forth in SEQ ID NO:1, 2, or 3. Such probes can then be used to specifically detect which allelic variant of a polymorphic region of a 5-LO gene is present in a patient. The polymorphic region can be located in the 5' upstream regulatory element, exon, or intron sequences of a 5-LO gene.

For example, preferred probes of the invention comprise a sequence listed in Table 2 or a complement thereof, wherein the bold nucleotides represent the location of the nucleotide polymorphism.

Particularly, preferred probes of the invention have a number of nucleotides sufficient to allow specific hybridization to the target nucleotide sequence. Where the target nucleotide sequence is present in a large fragment of DNA, such as a genomic DNA fragment of several tens or hundreds of kilobases, the size of the probe may have to be longer to provide sufficiently specific hybridization, as compared to a probe which is used to detect a target sequence which is present in a shorter fragment of DNA. For example, in some diagnostic methods, a portion of a 5-LO gene may first be amplified and thus isolated from the rest of the chromosomal DNA and then hybridized to a probe. In such a situation, a shorter probe will likely provide sufficient specificity of hybridization. For example, a probe having a nucleotide sequence of about 10 nucleotides may be sufficient.

In preferred embodiments, the probe or primer further comprises a label attached thereto, which, *e.g.,* is capable of being detected, *e.g.* the label group is selected from amongst radioisotopes, fluorescent compounds, enzymes, and enzyme co-factors.

In a preferred embodiment of the invention, the isolated nucleic acid, which is used, *e.g.,* as a probe or a primer, is modified, so as to be more stable than naturally occurring nucleotides. Exemplary nucleic acid molecules which are modified include phosphoramidate, phosphothioate and methylphosphonate analogs of DNA (see also U.S. Patent Numbers 5,176,996; 5,264,564; and 5,256,775).

The nucleic acids of the invention can also be modified at the base moiety, sugar moiety, or phosphate backbone, for example, to improve stability of the molecule. The nucleic acids, *e.g.,* probes or primers, may include other appended groups such as peptides (*e.g.,* for targeting host cell receptors *in vivo*), or agents facilitating transport across the cell membrane (see, *e.g.,* Letsinger *et al.*, 1989, *Proc. Natl. Acad. Sci. U.S.A.* 86:6553-6556; Lemaitre *et al.*, 1987, *Proc. Natl. Acad. Sci.* 84:648-652; PCT Publication No. WO88/09810, published December 15, 1988), hybridization-triggered cleavage agents. (See, *e.g.,* Krol *et al.*, 1988, *BioTechniques* 6:958-976) or intercalating agents. (See, *e.g.,* Zon, 1988, *Pharm. Res.* 5:539-549). To this end, the nucleic acid of the invention may be conjugated to another molecule, *e.g.,* a peptide, hybridization triggered cross-linking agent, transport agent, hybridization-triggered cleavage agent, etc.

The isolated nucleic acid comprising a 5-LO intronic sequence may comprise at least one modified base moiety which is selected from the group including but not limited to 5-fluorouracil, 5-bromouracil, 5-chlorouracil, 5-iodouracil, hypoxanthine, xantine, 4-acetylcytidine, 5-(carboxyhydroxymethyl) uracil, 5-carboxymethylaminomethyl-2-thiouridine, 5-carboxymethylaminomethyluracil, dihydrouracil, beta-D-galactosylqueosine, inosine, N6-isopentenyladenine, 1-methylguanine, 1-methylinosine, 2,2-dimethylguanine, 2-methyladenine, 2-methylguanine, 3-methylcytidine, 5-methylcytidine, N6-adenine, 7-methylguanine, 5-methylaminomethyluracil, 5-methoxyaminomethyl-2-thiouracil, beta-D-mannosylqueosine, 5'-methoxycarboxymethyluracil, 5-methoxyuracil, 2-methylthio-N6-isopentenyladenine, uracil-5-oxyacetic acid (v), wybutoxosine, pseudouracil, queosine, 2-thiocytidine, 5-methyl-2-thiouracil, 2-thiouracil, 4-thiouracil, 5-methyluracil, uracil-5-oxyacetic acid methylester, uracil-5-oxyacetic acid (v), 5-methyl-2-thiouracil, 3-(3-amino-3-N-2-carboxypropyl) uracil, (acp3)w, and 2,6-diaminopurine.

The isolated nucleic acid may also comprise at least one modified sugar moiety selected from the group including but not limited to arabinose, 2-fluoroarabinose, xylulose, and hexose.

In yet another embodiment, the nucleic acid comprises at least one modified phosphate backbone selected from the group consisting of a phosphorothioate, a phosphorodithioate, a phosphoramidothioate, a phosphoramidate, a phosphordiamidate, a methylphosphonate, an alkyl phosphotriester, and a formacetal or analog thereof.

In yet a further embodiment, the nucleic acid is an α-anomeric oligonucleotide. An α-anomeric oligonucleotide forms specific double-stranded hybrids with complementary RNA in which, contrary to the usual β-units, the strands run parallel to each other (Gautier *et al.*, 1987, *Nucl. Acids Res.* 15:6625-6641). The oligonucleotide is a 2'-0-methylribonucleotide (Inoue *et al.*, 1987, *Nucl. Acids Res.* 15:6131-6148), or a chimeric RNA-DNA analogue (Inoue *et al.*, 1987, *FEBS Lett.* 215:327-330).

Any nucleic acid fragment of the invention can be prepared according to methods well known in the art and described, *e.g.,* in Sambrook, J. Fritsch, E.F., and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual,* Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y. For example, discrete fragments of the DNA can be prepared and cloned using restriction enzymes. Alternatively, discrete fragments can be prepared using the Polymerase Chain Reaction (PCR) using primers having an appropriate sequence.

Oligonucleotides of the invention may be synthesized by standard methods known in the art, *e.g.* by use of an automated DNA synthesizer (such as are commercially available from Biosearch, Applied Biosystems, etc.). As examples, phosphorothioate oligonucleotides may be synthesized by the method of Stein *et al.* (1988, *Nucl. Acids Res.* 16:3209), methylphosphonate oligonucleotides can be prepared by use of controlled pore glass polymer supports (Sarin *et al.*, 1988, *Proc. Natl. Acad. Sci. U.S.A.* 85:7448-7451), etc.

The invention also provides vectors and plasmids comprising the nucleic acids of the invention. For example, in one embodiment, the invention provides a vector comprising at least a portion of a 5-LO gene comprising a polymorphic region. Thus, the invention provides vectors for expressing at least a portion of the newly identified allelic variants of the human 5-LO gene reference, as well as other allelic variants, comprising a nucleotide

sequence which is different from the nucleotide sequence disclosed in GI 187166, GI 8247778, or GI 450256. The allelic variants can be expressed in eukaryotic cells, *e.g.*, cells of a patient, or in prokaryotic cells.

In one embodiment, the vector comprising at least a portion of a 5-LO allele is introduced into a host cell, such that a protein encoded by the allele is synthesized. The 5-LO protein produced can be used, *e.g.*, for the production of antibodies, which can be used, *e.g.*, in methods for detecting mutant forms of 5-LO. Alternatively, the vector can be used for gene therapy, and be, *e.g.*, introduced into a patient to produce 5-LO protein. Host cells comprising a vector having at least a portion of a 5-LO gene are also within the scope of the invention.

Methods

The invention further provides predictive medicine methods, which are based, at least in part, on the discovery of 5-LO polymorphisms which are associated with specific physiological states and/or diseases or disorders.

For example, information obtained using the diagnostic assays described herein is useful for determining that a patient suffering from an inflammatory or allergy disease or disorder, *e.g.*, asthma, has a more or less severe disease phenotype, *e.g.*, a more or less severe asthma phenotype. Alternatively, the information can be used prognostically for predicting whether a patient will be responsive to treatment of an inflammatory or allergy disease or disorder, including, but not limited to, asthma, with a 5-LO inhibitor or other related agent. Based on the prognostic information, a health care provider can recommend a regimen (*e.g.*, diet or exercise) or therapeutic protocol, (*e.g.*, administration of a 5-LO inhibitor) useful for preventing the particular disease or disorder, or symptoms of the disease or disorder in the patient *e.g.*, asthma.

In addition, knowledge of the identity of a particular 5-LO allele in an individual (the 5-LO genetic profile, including the presence or absence of the polymorphisms of the haplotype as described herein), alone or in conjunction with information on other genetic defects contributing to the same disease (the genetic profile of the particular disease, *e.g.*, asthma) allows customization of therapy for a particular disease to the individual's genetic

profile. For example, an individual's 5-LO genetic profile or the genetic profile of a disease or condition associated with a specific allele of a 5-LO polymorphic region, can enable a health care provider: 1) to more effectively prescribe a drug that will address the molecular basis of the disease or condition (*e.g.*, a 5-LO inhibitor as described in U.S. Patent Number 5,703,093, 5,750,565, or 6,025,384); and 2) to better determine the appropriate dosage of a particular drug. For example, the expression level of 5-LO proteins, alone or in conjunction with the expression level of other genes, known to contribute to the same disease, can be measured in many patients at various stages of the disease to generate a transcriptional or expression profile of the disease. Expression patterns of individual patients can then be compared to the expression profile of the disease to determine the appropriate drug and dose to administer to the patient.

The ability to target populations expected to show the highest clinical benefit, *e.g.*, subgroups of asthmatic populations which respond or do not respond to specific therapies, *e.g.*, treatment with 5-LO inhibitors, or subgroups of asthmatic populations with more or less severe asthma phenotypes, based on the 5-LO or disease genetic profile, can enable: 1) the repositioning of marketed drugs with disappointing market results; 2) the rescue of drug candidates whose clinical development has been discontinued as a result of safety or efficacy limitations, which are patient subgroup-specific; and 3) an accelerated and less costly development for drug candidates and more optimal drug labeling (*e.g.*, since the use of 5-LO as a marker is useful for optimizing effective dose).

These and other methods are described in further detail in the following sections.

A. Prognostic and Diagnostic Assays

The present methods provide means for determining if a patient has a disease, condition or disorder that is associated a specific 5-LO allele, *e.g.*, aberrant, *e.g.*, decreased, eosinophil levels, or the severity of an inflammatory disease or disorder phenotype, *e.g.*, the severity of an asthma phenotype among an asthmatic population. The present methods also provide means for predicting or determining if a patient will be more or less responsive to treatment with a 5-LO inhibitor, based on determination of a specific 5-LO allele.

The present invention provides methods for determining the molecular structure of a 5-LO gene, such as a human 5-LO gene, or a portion thereof. For example, the present invention provides methods for determining the presence or absence of the haplotype described herein. In one embodiment, determining the molecular structure of at least a portion of a 5-LO gene comprises determining the identity of the allelic variant of at least one polymorphic region of a 5-LO gene (determining the presence or absence of one or more of the allelic variants, or their complements, of SEQ ID NOs.:4-6, thereby determining the presence of the haplotype as described herein). A polymorphic region of a 5-LO gene can be located in an exon, an intron, at an intron/exon border, or, preferably, in the 5' upstream regulatory element of the 5-LO gene.

The invention provides methods for determining whether a patient has a specific disease or disorder phenotype, *e.g.*, asthma phenotype, associated with a specific allelic variant of a polymorphic region of a 5-LO gene. Such disease phenotypes are associated with aberrant 5-LO activity, *e.g.*, aberrant 5-LO expression, or increased or decreased eosinophil counts in the patient. Aberrant 5-LO protein level can result from aberrant transcription or post transcriptional regulation. Thus, allelic differences in specific regions of a 5-LO gene result in differences of 5-LO protein due to differences in regulation of expression. In particular, some of the identified polymorphisms in the human 5-LO gene are associated with differences in the level of transcription, RNA maturation, splicing, or translation of the 5-LO gene or transcription product. For example, the presence of a variant Sp1 binding site, which is associated with the haplotype of the invention, result in reduced levels of 5-LO expression. Therefore, presence of any of the polymorphisms of the haplotype predicts a reduced levels of 5-LO expression.

Analysis of one or more 5-LO polymorphic regions in a patient, including the identities of the polymorphisms included in the haplotype of the present invention can be useful for predicting or determining whether a patient has more or less severe inflammatory disease phenotype, *e.g.*, a more or less severe asthma phenotype, associated with increased or decreased eosinophil levels resulting in more or less severe symptoms of the inflammatory disease or disorder.

In preferred embodiments, the methods of the invention can be characterized as comprising detecting, in a sample of cells or nucleic acid from the patient, the presence or absence of a specific allelic variant of one or more polymorphic regions of a 5-LO gene. The allelic differences can be: (i) a difference in the identity of at least one nucleotide or (ii) a difference in the number of nucleotides, which difference can be a single nucleotide or several nucleotides. The invention also provides methods for detecting differences in 5-LO genes such as chromosomal rearrangements, *e.g.,* chromosomal dislocation. The invention can also be used in prenatal diagnostics.

A preferred detection method is allele specific hybridization using probes overlapping the polymorphic site and having about 5, 10, 20, 25, or 30 nucleotides around the polymorphic region. Examples of probes for detecting specific allelic variants of the polymorphic region located in the 5' upstream regulatory element of the 5-LO gene are probes comprising a nucleotide sequence set forth in any of SEQ ID NOs.:64, 65, and 66 or the complement thereof. In a preferred embodiment of the invention, several probes capable of hybridizing specifically to allelic variants are attached to a solid phase support, *e.g.,* a "chip". Oligonucleotides can be bound to a solid support by a variety of processes, including lithography. For example a chip can hold up to 250,000 oligonucleotides (GeneChip, Affymetrix). Mutation detection analysis using these chips comprising oligonucleotides, also termed "DNA probe arrays" is described *e.g.,* in Cronin *et al.* (1996) Human Mutation 7:244. In one embodiment, a chip comprises all the allelic variants of at least one polymorphic region of a gene. The solid phase support is then contacted with a test nucleic acid and hybridization to the specific probes is detected. Accordingly, the identity of numerous allelic variants of one or more genes can be identified in a simple hybridization experiment. For example, the identity of the allelic variant of the nucleotide polymorphism in the 5' upstream regulatory element can be determined in a single hybridization experiment.

In other detection methods, it is necessary to first amplify at least a portion of a 5-LO gene prior to identifying the allelic variant. Amplification can be performed, *e.g.,* by PCR and/or LCR (see Wu and Wallace, (1989) *Genomics* 4:560), according to methods known in the art. In one embodiment, genomic DNA of a cell is exposed to two PCR primers and amplification for a number of cycles sufficient to produce the required amount

of amplified DNA. In preferred embodiments, the primers are located between 150 and 350 base pairs apart. Preferred primers, such as primers for amplifying each of the exons of the human 5-LO gene, are listed in Table 3.

Alternative amplification methods include: self sustained sequence replication (Guatelli, J.C. *et al.*, 1990, *Proc. Natl. Acad. Sci. USA* 87:1874-1878), transcriptional amplification system (Kwoh, D.Y. *et al.*, 1989, *Proc. Natl. Acad. Sci. USA* 86:1173-1177), Q-Beta Replicase (Lizardi, P.M. *et al.*, 1988, Bio/Technology 6:1197), and self-sustained sequence replication (*Guatelli et al.,* (1989) *Proc. Nat. Acad. Sci.* 87:1874), and nucleic acid based sequence amplification (NABSA), or any other nucleic acid amplification method, followed by the detection of the amplified molecules using techniques well known to those of skill in the art. These detection schemes are especially useful for the detection of nucleic acid molecules if such molecules are present in very low numbers.

In one embodiment, any of a variety of sequencing reactions known in the art can be used to directly sequence at least a portion of a 5-LO gene and detect allelic variants, *e.g.,* mutations, by comparing the sequence of the sample sequence with the corresponding wild-type (control) sequence. Exemplary sequencing reactions include those based on techniques developed by Maxam and Gilbert (*Proc. Natl Acad Sci USA* (1977) 74:560) or Sanger (Sanger *et al* (1977) *Proc. Nat. Acad. Sci* 74:5463). It is also contemplated that any of a variety of automated sequencing procedures may be utilized when performing the subject assays (*Biotechniques* (1995) 19:448), including sequencing by mass spectrometry (see, for example, U.S. Patent Number 5,547,835 and international patent application Publication Number WO 94/16101, entitled *DNA Sequencing by Mass Spectrometry* by H. Köster; U.S. Patent Number 5,547,835 and international patent application Publication Number WO 94/21822 entitled "DNA Sequencing by Mass Spectrometry Via Exonuclease Degradation" by H. Köster), and U.S Patent Number 5,605,798 and International Patent Application No. PCT/US96/03651 entitled *DNA Diagnostics Based on Mass Spectrometry* by H. Köster;. Cohen *et al.* (1996) *Adv Chromatogr* 36:127-162; and Griffin *et al.* (1993) *Appl Biochem Biotechnol* 38:147-159). It will be evident to one skilled in the art that, for certain embodiments, the occurrence of only one, two or three of the nucleic acid bases need be

determined in the sequencing reaction. For instance, A-track or the like, *e.g.,* where only one nucleotide is detected, can be carried out.

Yet other sequencing methods are disclosed, *e.g.,* in U.S. Patent Number 5,580,732 entitled "Method of DNA sequencing employing a mixed DNA-polymer chain probe" and U.S. Patent Number 5,571,676 entitled "Method for mismatch-directed *in vitro* DNA sequencing".

In some cases, the presence of a specific allele of a 5-LO gene in DNA from a patient can be shown by restriction enzyme analysis. For example, a specific nucleotide polymorphism can result in a nucleotide sequence comprising a restriction site which is absent from the nucleotide sequence of another allelic variant.

In a further embodiment, protection from cleavage agents (such as a nuclease, hydroxylamine or osmium tetroxide and with piperidine) can be used to detect mismatched bases in RNA/RNA DNA/DNA, or RNA/DNA heteroduplexes (Myers, *et al.* (1985) *Science* 230:1242). In general, the technique of "mismatch cleavage" starts by providing heteroduplexes formed by hybridizing a control nucleic acid, which is optionally labeled, *e.g.,* RNA or DNA, comprising a nucleotide sequence of a 5-LO allelic variant with a sample nucleic acid, *e.g.,* RNA or DNA, obtained from a tissue sample. The double-stranded duplexes are treated with an agent which cleaves single-stranded regions of the duplex such as duplexes formed based on basepair mismatches between the control and sample strands. For instance, RNA/DNA duplexes can be treated with RNase and DNA/DNA hybrids treated with S1 nuclease to enzymatically digest the mismatched regions. In other embodiments, either DNA/DNA or RNA/DNA duplexes can be treated with hydroxylamine or osmium tetroxide and with piperidine in order to digest mismatched regions. After digestion of the mismatched regions, the resulting material is then separated by size on denaturing polyacrylamide gels to determine whether the control and sample nucleic acids have an identical nucleotide sequence or in which nucleotides they are different. See, for example, Cotton *et al* (1988) *Proc. Natl Acad Sci USA* 85:4397; Saleeba *et al* (1992) *Methods Enzymol.* 217:286-295. In a preferred embodiment, the control or sample nucleic acid is labeled for detection.

In another embodiment, an allelic variant can be identified by denaturing high-

performance liquid chromatography (DHPLC) (Oefner and Underhill, (1995) *Am. J. Human Gen.* 57:Suppl. A266). DHPLC uses reverse-phase ion-pairing chromatography to detect the heteroduplexes that are generated during amplification of PCR fragments from individuals who are heterozygous at a particular nucleotide locus within that fragment (Oefner and Underhill (1995) *Am. J. Human Gen.* 57:Suppl. A266). In general, PCR products are produced using PCR primers flanking the DNA of interest. DHPLC analysis is carried out and the resulting chromatograms are analyzed to identify base pair alterations or deletions based on specific chromatographic profiles (see O'Donovan *et al.* (1998) *Genomics* 52:44-49).

In other embodiments, alterations in electrophoretic mobility is used to identify the type of 5-LO allelic variant. For example, single strand conformation polymorphism (SSCP) may be used to detect differences in electrophoretic mobility between mutant and wild type nucleic acids (Orita *et al.* (1989) *Proc Natl. Acad. Sci USA* 86:2766, see also Cotton (1993) *Mutat Res* 285:125-144; and Hayashi (1992) *Genet Anal Tech Appl* 9:73-79). Single-stranded DNA fragments of sample and control nucleic acids are denatured and allowed to renature. The secondary structure of single-stranded nucleic acids varies according to sequence, the resulting alteration in electrophoretic mobility enables the detection of even a single base change. The DNA fragments may be labeled or detected with labeled probes. The sensitivity of the assay may be enhanced by using RNA (rather than DNA), in which the secondary structure is more sensitive to a change in sequence. In another preferred embodiment, the subject method utilizes heteroduplex analysis to separate double stranded heteroduplex molecules on the basis of changes in electrophoretic mobility (Keen *et al.* (1991) *Trends Genet* 7:5).

In yet another embodiment, the identity of an allelic variant of a polymorphic region is obtained by analyzing the movement of a nucleic acid comprising the polymorphic region in polyacrylamide gels containing a gradient of denaturant is assayed using denaturing gradient gel electrophoresis (DGGE) (Myers *et al* (1985) *Nature* 313:495). When DGGE is used as the method of analysis, DNA will be modified to insure that it does not completely denature, for example by adding a GC clamp of approximately 40 bp of high-melting GC-rich DNA by PCR. In a further embodiment, a temperature gradient is used in place of a

denaturing agent gradient to identify differences in the mobility of control and sample DNA (Rosenbaum and Reissner (1987) *Biophys Chem* 265:1275).

Examples of techniques for detecting differences of at least one nucleotide between 2 nucleic acids include, but are not limited to, selective oligonucleotide hybridization, selective amplification, or selective primer extension. For example, oligonucleotide probes may be prepared in which the known polymorphic nucleotide is placed centrally (allele-specific probes) and then hybridized to target DNA under conditions which permit hybridization only if a perfect match is found (Saiki *et al.* (1986) *Nature* 324:163); Saiki *et al* (1989) *Proc. Natl Acad. Sci USA* 86:6230; and Wallace *et al.* (1979) *Nucl. Acids Res.* 6:3543). Such allele specific oligonucleotide hybridization techniques may be used for the simultaneous detection of several nucleotide changes in different polylmorphic regions of 5-LO. For example, oligonucleotides having nucleotide sequences of specific allelic variants are attached to a hybridizing membrane and this membrane is then hybridized with labeled sample nucleic acid. Analysis of the hybridization signal will then reveal the identity of the nucleotides of the sample nucleic acid.

Alternatively, allele specific amplification technology which depends on selective PCR amplification may be used in conjunction with the instant invention. Oligonucleotides used as primers for specific amplification may carry the allelic variant of interest in the center of the molecule (so that amplification depends on differential hybridization) (Gibbs *et al* (1989) *Nucleic Acids Res.* 17:2437-2448) or at the extreme 3' end of one primer where, under appropriate conditions, mismatch can prevent, or reduce polymerase extension (Prossner (1993) *Tibtech* 11:238; Newton *et al.* (1989) *Nucl. Acids Res.* 17:2503). This technique is also termed "PROBE" for Probe Oligo Base Extension. In addition it may be desirable to introduce a novel restriction site in the region of the mutation to create cleavage-based detection (Gasparini *et al* (1992) *Mol. Cell Probes* 6:1).

In another embodiment, identification of the allelic variant is carried out using an oligonucleotide ligation assay (OLA), as described, *e.g.,* in U.S. Patent Number 4,998,617 and in Landegren, U. *et al.*, (1988) *Science* 241:1077-1080. The OLA protocol uses two oligonucleotides which are designed to be capable of hybridizing to abutting sequences of a single strand of a target. One of the oligonucleotides is linked to a separation marker, *e.g.*,

biotinylated, and the other is detectably labeled. If the precise complementary sequence is found in a target molecule, the oligonucleotides will hybridize such that their termini abut, and create a ligation substrate. Ligation then permits the labeled oligonucleotide to be recovered using avidin, or another biotin ligand. Nickerson, D. A. *et al.* have described a nucleic acid detection assay that combines attributes of PCR and OLA (Nickerson, D. A. *et al.*, (1990) *Proc. Natl. Acad. Sci. (U.S.A.)* 87:8923-8927. In this method, PCR is used to achieve the exponential amplification of target DNA, which is then detected using OLA.

Several techniques based on this OLA method have been developed and can be used to detect specific allelic variants of a polymorphic region of a 5-LO gene. For example, U.S. Patent Number 5,593,826 discloses an OLA using an oligonucleotide having 3'-amino group and a 5'-phosphorylated oligonucleotide to form a conjugate having a phosphoramidate linkage. In another variation of OLA described in Tobe *et al.* ((1996) *Nucleic Acids Res* 24: 3728), OLA combined with PCR permits typing of two alleles in a single microtiter well. By marking each of the allele-specific primers with a unique hapten, i.e., digoxigenin and fluorescein, each OLA reaction can be detected by using hapten specific antibodies that are labeled with different enzyme reporters, alkaline phosphatase or horseradish peroxidase. This system permits the detection of the two alleles using a high throughput format that leads to the production of two different colors.

The invention further provides methods for detecting single nucleotide polymorphisms in a 5-LO gene. Because single nucleotide polymorphisms constitute sites of variation flanked by regions of invariant sequence, their analysis requires no more than the determination of the identity of the single nucleotide present at the site of variation and it is unnecessary to determine a complete gene sequence for each patient. Several methods have been developed to facilitate the analysis of such single nucleotide polymorphisms.

In one embodiment, the single base polymorphism can be detected by using a specialized exonuclease-resistant nucleotide, as disclosed, *e.g.,* in Mundy, C. R. (U.S. Patent Number 4,656,127). According to the method, a primer complementary to the allelic sequence immediately 3' to the polymorphic site is permitted to hybridize to a target molecule obtained from a particular animal or human. If the polymorphic site on the target molecule contains a nucleotide that is complementary to the particular exonuclease-resistant

nucleotide derivative present, then that derivative will be incorporated onto the end of the hybridized primer. Such incorporation renders the primer resistant to exonuclease, and thereby permits its detection. Since the identity of the exonuclease-resistant derivative of the sample is known, a finding that the primer has become resistant to exonucleases reveals that the nucleotide present in the polymorphic site of the target molecule was complementary to that of the nucleotide derivative used in the reaction. This method has the advantage that it does not require the determination of large amounts of extraneous sequence data.

In another embodiment of the invention, a solution-based method is used for determining the identity of the nucleotide of a polymorphic site. Cohen, D. *et al.* (French Patent 2,650,840; PCT Appln. No. WO91/02087). As in the Mundy method of U.S. Patent Number 4,656,127, a primer is employed that is complementary to allelic sequences immediately 3' to a polymorphic site. The method determines the identity of the nucleotide of that site using labeled dideoxynucleotide derivatives, which, if complementary to the nucleotide of the polymorphic site will become incorporated onto the terminus of the primer.

An alternative method, known as Genetic Bit Analysis or GBA™ is described by Goelet, P. *et al.* (PCT Appln. No. 92/15712). The method of Goelet, P. *et al.* uses mixtures of labeled terminators and a primer that is complementary to the sequence 3' to a polymorphic site. The labeled terminator that is incorporated is thus determined by, and complementary to, the nucleotide present in the polymorphic site of the target molecule being evaluated. In contrast to the method of Cohen *et al.* (French Patent 2,650,840; PCT Appln. No. WO91/02087) the method of Goelet, P. *et al.* is preferably a heterogeneous phase assay, in which the primer or the target molecule is immobilized to a solid phase.

Recently, several primer-guided nucleotide incorporation procedures for assaying polymorphic sites in DNA have been described (Komher, J. S. *et al.*, *Nucl. Acids. Res.* 17:7779-7784 (1989); Sokolov, B. P., Nucl. Acids Res. 18:3671 (1990); Syvanen, A. -C., *et al.*, *Genomics* 8:684-692 (1990); Kuppuswamy, M. N. *et al.*, *Proc. Natl. Acad. Sci. (U.S.A.)* 88:1143-1147 (1991); Prezant, T. R. *et al.*, *Hum. Mutat.* 1:159-164 (1992); Ugozzoli, L. *et al.*, GATA 9:107-112 (1992); Nyren, P. *et al.*, *Anal. Biochem.* 208:171-175 (1993)). These methods differ from GBA™ in that they all rely on the incorporation of labeled deoxynucleotides to discriminate between bases at a polymorphic site. In such a format,

since the signal is proportional to the number of deoxynucleotides incorporated, polymorphisms that occur in runs of the same nucleotide can result in signals that are proportional to the length of the run (Syvanen, A. -C., *et al.*, *Amer. J. Hum. Genet.* 52:46-59 (1993)).

If a polymorphic region is located in an exon, either in a coding or non-coding portion of the gene, the identity of the allelic variant can be determined by determining the molecular structure of the mRNA, pre-mRNA, or cDNA. The molecular structure can be determined using any of the above described methods for determining the molecular structure of the genomic DNA, *e.g.*, see Example 1.

The methods described herein may be performed, for example, by utilizing pre-packaged diagnostic kits, such as those described above, comprising at least one probe or primer nucleic acid described herein, which may be conveniently used, *e.g.*, to determine whether a patient has a specific disease phenotype associated with a specific 5-LO allelic variant.

Sample nucleic acid to be analyzed by any of the above-described diagnostic and prognostic methods can be obtained from any cell type or tissue of a patient. For example, a patient's bodily fluid (*e.g.*, blood) can be obtained by known techniques (*e.g.* venipuncture). Alternatively, nucleic acid tests can be performed on dry samples (*e.g.* hair or skin). Fetal nucleic acid samples can be obtained from maternal blood as described in International Patent Application No. WO91/07660 to Bianchi. Alternatively, amniocytes or chorionic villi may be obtained for performing prenatal testing.

Diagnostic procedures may also be performed *in situ* directly upon tissue sections (fixed and/or frozen) of patient tissue obtained from biopsies or resections, such that no nucleic acid purification is necessary. Nucleic acid reagents may be used as probes and/or primers for such *in situ* procedures (see, for example, Nuovo, G.J., 1992, PCR *in situ* hybridization: protocols and applications, Raven Press, NY).

In addition to methods which focus primarily on the detection of one nucleic acid sequence, profiles may also be assessed in such detection schemes. Fingerprint profiles may be generated, for example, by utilizing a differential display procedure, Northern analysis and/or RT-PCR.

B. Pharmacogenomics

Knowledge of the identity of the allele of one or more 5-LO gene polymorphic regions in a patient (the 5-LO genetic profile), alone or in conjunction with information on other genetic defects contributing to the same disease (the genetic profile of the particular disease, including, for example, the determination of the presence of the haplotype described herein) also allows a customization of the therapy for a particular disease to the patient's genetic profile.

It has been determined that the presence of a variant Sp1 binding site in the promoter region of the 5-LO gene results in reduced responsiveness to treatment with particular therapies, including treatment with 5-LO inhibitors, or other agents which antagonize 5-LO in asthmatic patients (U.S. Patent Number 6,090,547). Therefore, based on the existence of the haplotype as described herein which is in linkage disequilibrium with any one of three variant Sp1 binding site, the detection in an asthmatic patient of an allele present in at least one of the polymorphic sites included in the haplotype of the present inventioncan be used to predict whether or not that the patient has reduced responsiveness to treatment with a specific treatment, *e.g.,* treatment with a 5-LO inhibitor. The methods of the invention further provide methods for identifying a candidate for treatment with a 5-LO inhibitor by the detection in a patient of an allele present in at least one of the polymorphic sites included in the haplotype of the present invention, or a complement thereof.

For example, patients having a specific allele of a 5-LO gene may or may not exhibit symptoms of a particular disease or be predisposed to developing symptoms of a particular disease. Further, if those patients are symptomatic, they may or may not respond well to a certain drug, *e.g.,* a specific 5-LO therapeutic, such as a 5-LO inhibitor, but may respond to another. Thus, generation of a 5-LO genetic profile (*e.g.,* categorization of alterations in 5-LO genes which are associated with the development of a particular disease, including determination of the presence of the haplotype described herein), from a population of patients, who are symptomatic for a disease or condition that is caused by or contributed to by a defective and/or aberrantly expressed 5-LO gene and/or protein (a 5-LO genetic population profile) and comparison of an individual's 5-LO profile to the population profile,

permits the selection or design of drugs that are expected to be safe and efficacious for a particular patient or patient population (*i.e.*, a group of patients having the same genetic alteration).

For example, a 5-LO population profile can be performed by determining the 5-LO profile, *e.g.*, the identity of 5-LO alleles, in particular the identity of 5-LO alleles included in the haplotype as described herein, in a patient population having a disease, which is associated with one or more specific alleles of 5-LO polymorphic regions. Optionally, the 5-LO population profile can further include information relating to the response of the population to a 5-LO therapeutic, using any of a variety of methods, including, monitoring: 1) the severity of symptoms associated with the 5-LO related disease, 2) 5-LO gene expression level, 3) 5-LO mRNA level, 4) 5-LO protein level, 5) eosinophil level, and/or 6) leukotriene level, and dividing or categorizing the population based on particular 5-LO alleles. The 5-LO genetic population profile can also, optionally, indicate those particular 5-LO alleles which are present in patients that are either responsive or non-responsive to a particular therapeutic, *e.g.*, a 5-LO inhibitor. This information or population profile, is then useful for predicting which individuals should respond to particular drugs, based on their individual 5-LO profile.

In a preferred embodiment, the 5-LO profile is a transcriptional or expression level profile and is comprised of determining the expression level of 5-LO proteins, alone or in conjunction with the expression level of other genes known to contribute to the same disease at various stages of the disease.

Pharmacogenomic studies can also be performed using transgenic animals. For example, one can produce transgenic mice, *e.g.*, as described herein, which contain a specific allelic variant of a 5-LO gene. These mice can be created, *e.g.*, by replacing their wild-type 5-LO gene with an allele of the human 5-LO gene. The response of these mice to specific 5-LO therapeutics can then be determined.

C. Monitoring Effects of 5-LO Therapeutics During Clinical Trials

The present invention provides a method for monitoring the effectiveness of treatment of a patient with an agent (*e.g.*, an agonist, antagonist, peptidomimetic, protein,

peptide, nucleic acid, small molecule, or other drug candidate identified, *e.g.,* by the screening assays described herein) comprising the steps of (i) obtaining a preadministration sample from a patient prior to administration of the agent; (ii) detecting the level of expression or activity of a 5-LO protein, mRNA or gene in the preadministration sample; (iii) obtaining one or more post-administration samples from the patient; (iv) detecting the level of expression or activity of the 5-LO protein, mRNA or gene in the post-administration samples; (v) comparing the level of expression or activity of the 5-LO protein, mRNA, or gene in the preadministration sample with those of the 5-LO protein, mRNA, or gene in the post administration sample or samples; and (vi) altering the administration of the agent to the patient accordingly. For example, increased administration of the agent may be desirable to increase the expression or activity of 5-LO to higher levels than detected, *i.e.,* to increase the effectiveness of the agent. Alternatively, decreased administration of the agent may be desirable to decrease expression or activity of 5-LO to lower levels than detected, *i.e.,* to decrease the effectiveness of the agent.

Cells of a patient may also be obtained before and after administration of a 5-LO therapeutic to detect the level of expression of genes other than 5-LO, to verify that the 5-LO therapeutic does not increase or decrease the expression of genes which could be deleterious. This can be done, *e.g.,* by using the method of transcriptional profiling. Thus, mRNA from cells exposed *in vivo* to a 5-LO therapeutic and mRNA from the same type of cells that were not exposed to the 5-LO therapeutic could be reverse transcribed and hybridized to a chip containing DNA from numerous genes, to thereby compare the expression of genes in cells treated and not treated with a 5-LO therapeutic. If, for example a 5-LO therapeutic turns on the expression of a proto-oncogene in an individual, use of this particular 5-LO therapeutic may be undesirable.

D. Methods of Treatment

The present invention provides for both prophylactic and therapeutic methods of treating a patient having a disorder associated with specific 5-LO alleles and/or aberrant 5-LO expression or activity, *e.g.,* inflammatory or allergy diseases or disorders, such as asthma.

i) Prophylactic Methods

In one aspect, the invention provides a method for preventing in a patient, a disease or condition associated with a specific 5-LO allele such as an inflammatory or allergy disease or disorder, *e.g.,* asthma, and medical conditions resulting therefrom, by administering to the patient an agent which counteracts the unfavorable biological effect of the specific 5-LO allele. Subjects at risk for such a disease can be identified by a diagnostic or prognostic assay, *e.g.,* as described herein. Administration of a prophylactic agent can occur prior to the manifestation of symptoms associated with specific 5-LO alleles, such that a disease or disorder is prevented or, alternatively, delayed in its progression. Depending on the identity of the 5-LO allele in a patient, a compound that counteracts the effect of this allele is administered. The compound can be a compound modulating the activity of 5-LO, *e.g.,* a 5-LO inhibitor. The treatment can also be a specific diet, or environmental alteration. In particular, the treatment can be undertaken prophylactically, before any other symptoms are present. Such a prophylactic treatment could thus prevent the development of an aberrant inflammatory response, *e.g.,* asthma. The prophylactic methods are similar to therapeutic methods of the present invention and are further discussed in the following subsections.

(ii) Therapeutic Methods

The invention further provides methods of treating patients having a specific disease or disorder phenotype associated with a specific allelic variant of a polymorphic region of a 5-LO gene, *e.g.,* a more moderate or more severe phenotype. Preferred diseases or disorders include those associated with aberrant leukotriene synthesis, eosinophil levels, and 5-LO expression, and disorders resulting therefrom .

In one embodiment, the method comprises (a) determining the identity of one or more allelic variants of the invention; and (b) administering to the patient a compound that compensates for the effect of the specific allelic variant. The polymorphic region can be localized at any location of the gene, *e.g.,* in a regulatory element (*e.g.,* in a 5' upstream regulatory element), in an exon, (*e.g.,* coding region of an exon), in an intron, or at an exon/intron border. Thus, depending on the site of the polymorphism in the 5-LO gene, a patient having a specific variant of the polymorphic region which is associated with a specific disease or condition, can be treated with compounds which specifically compensate for the effect of the allelic variant.

In a preferred embodiment, the identity of any one of the polymorphisms included in the haplotype is determined. In another preferred embodiment, the identity of one or more of the following nucleotides of a 5-LO gene of a patient is determined: adenine at residue 1000 of the reference sequence GI 187166, a deletion at residues 472-477 (GTTAAA) (SEQ ID NO:62) of the reference sequence GI 187166, and/or adenine at residue 559 of the reference sequence GI 187166 may be determined.

In another preferred embodiment, the identity of at least one of the nucleotides listed above is determined, in combination with one or more of the known nucleotide polymorphisms from Table 1, *i.e.,* adenine at residue 94 of the reference sequence GI 187166, adenine at residue 137 of the reference sequence GI 187166, and/or variant Sp1 binding site *e.g.,* sites containing three, four, or six Sp1 motifs, may be determined. In a particularly preferred embodiment, it is determined that a patient has one or more of the following: an adenine at residue 1000 of the reference sequence GI 187166, a deletion of residues 472-477 (GTTAAA) (SEQ ID NO:62) of the reference sequence GI 187166, an adenine at residue 559 of the reference sequence GI 187166, in combination with one or more of the following previously known polymorphisms, preferably an adenine at residue 84 of GI 187166, and/or an adenine at residue 137 of GI 187166, or a variant Sp1 binding site

If a patient having an inflammatory or allergy disease or disorder, *e.g.,* asthma, has one or more of the polymorphisms of the invention (as set forth in Table 1), or one or more of the polymorphisms of the invention in combination with one or more of the previously known polymorphisms listed in Table 1 or a variant Sp1 binding site allele, that patient has a

decreased risk of having high eosinophil levels, has a more moderate inflammatory disease phenotype, has a decreased risk of having increased 5-LO expression, and can be identified as having decreased responsiveness to a 5-LO inhibitor, as compared to a patient without one of the aforementioned polymorphisms.

Similarly, if a patient having an inflammatory disease or disorder, *e.g.,* asthma does not have one or more of the polymorphisms of the invention (as set forth in Table 1), or one or more of the polymorphisms of the invention in combination with one or more of the previously known polymorphisms listed in Table 1 or a variant Sp1 binding site, that patient has an increased risk of having increased eosinophil levels, an increased risk of having a more severe inflammatory disease phenotype, an increased risk of having increased 5-LO expression, and can be identified as having increased responsiveness to a 5-LO inhibitor as compared to a patient with one of the aforementioned polymorphisms.

Generally, the allelic variant can be a mutant allele, *i.e.,* an allele which when present in one, or preferably two copies, in a patient results in a change in the phenotype of the patient. A mutation can be a substitution, deletion, and/or addition of at least one nucleotide relative to the wild-type allele (*i.e.,* the reference sequence). Depending on where the mutation is located in the 5-LO gene, the patient can be treated to specifically compensate for the mutation. For example, if the mutation is present in the coding region of the gene and results in a more active 5-LO protein, the patient can be treated, *e.g.,* by administration to the patient of a 5-LO inhibitor, such that the administration of the inhibitor prevents aberrant leukotriene synthesis by the 5-LO protein. Normal 5-LO protein can also be used to counteract or compensate for the endogenous mutated form of the 5-LO protein. Normal 5-LO protein can be directly delivered to the patient or indirectly by gene therapy wherein some cells in the patient are transformed or transfected with an expression construct encoding wild-type 5-LO protein. Nucleic acids encoding wild-type human 5-LO protein are set forth in SEQ ID NOs.:1, 2, and 3 (GI Accession Nos. 187166, 8247778 and 4502056).

Furthermore, depending on the site of the mutation in the 5-LO protein and the specific effect on its activity, specific treatments can be designed to compensate for that effect. The 5-LO protein is a lipoxygenase which interacts with arachidonic acid to synthesize leukotrienes (LTs), *e.g.,* LTA4, B4, C4, D4, and E4. The structure of the 5-LO

protein is further described, *e.g.*, in Silverman, *et al.* (1999) *Proc. Assoc. Am. Physicians* 111(6):525-36 and Matsumoto, *et al.* (1988) *supra*). Thus, if the mutation results in a 5-LO protein which causes as increased ability of 5-LO to synthesize leukotrienes, resulting in an accumulation of leukotrienes in the patient, a treatment can be designed which removes excess leukotrienes from the patient or inhibits 5-LO from synthesis of leukotrienes. In one embodiment, a compound which inhibits 5-LO synthesis of leukotrienes is administered to the patient. For example, a compound described in U.S. Patent Number 5,703,093, 5,750,565, or 6,025,384 can be administered to the patient.

Yet in another embodiment, the invention provides methods for treating a patient having a mutated 5-LO gene, in which the mutation is located in a regulatory region of the gene. Such a regulatory region can be localized in the 5' upstream regulatory element of the gene, in the 5' or 3' untranslated region of an exon, or in an intron. A mutation in a regulatory region can result in increased production of 5-LO protein, decreased production of 5-LO protein, or production of 5-LO having an aberrant tissue distribution. The effect of a mutation in a regulatory region upon the 5-LO protein can be determined, *e.g.*, by measuring the 5-LO protein level or mRNA level in cells having a 5-LO gene having this mutation and which, normally (*i.e.*, in the absence of the mutation) produce 5-LO protein. The effect of a mutation can also be determined *in vitro*. For example, if the mutation is in the 5' upstream regulatory element, a reporter construct can be constructed which comprises the mutated 5' upstream regulatory element linked to a reporter gene, the construct transfected into cells, and comparison of the level of expression of the reporter gene under the control of the mutated 5' upstream regulatory element and under the control of a wild-type 5' upstream regulatory element. Such experiments can also be carried out in mice transgenic for the mutated 5' upstream regulatory element. If the mutation is located in an intron, the effect of the mutation can be determined, *e.g.*, by producing transgenic animals in which the mutated 5-LO gene has been introduced and in which the wild-type gene may have been knocked out. Comparison of the level of expression of 5-LO in the mice transgenic for the mutant human 5-LO gene with mice transgenic for a wild-type human 5-LO gene will reveal whether the mutation results in increased, or decreased synthesis of the 5-LO protein and/or aberrant tissue distribution of 5-LO protein. Such analysis could also be performed in cultured cells,

in which the human mutant 5-LO gene is introduced and, *e.g.,* replaces the endogenous wild-type 5-LO gene in the cell. Thus, depending on the effect of the mutation in a regulatory region of a 5-LO gene, a specific treatment can be administered to a patient having such a mutation. Accordingly, if the mutation results in decreased 5-LO protein levels, the patient can be identified as being less responsive to treatment with a 5-LO inhibitor. For example, the presence of a non-wild-type Sp1 binding site allele or a polymorphism included in a haplotype as described herein, may indicate reduced levels of 5-LO expression, reduced eosinophil levels, a less severe asthma phenotype, and reduced responsiveness to 5-LO inhibition by a 5-LO inhibitor, *e.g.,* a 5-LO inhibitor described in U.S. Patent Number 5,703,093, 5,750,565 or 6,025,384.

A correlation between drug responses and specific alleles of 5-LO can be shown, for example, by clinical studies wherein the response to specific drugs of patients having different allelic variants of a polymorphic region of a 5-LO gene is compared. Such studies can also be performed using animal models, such as mice having various alleles of human 5-LO genes and in which, *e.g.,* the endogenous 5-LO has been inactivated such as by a knock-out mutation. Test drugs are then administered to the mice having different human 5-LO alleles and the response of the different mice to a specific compound is compared. Accordingly, the invention provides assays for identifying the drug which will be best suited for treating a specific disease or condition in a patient. For example, it will be possible to select drugs which will be devoid of toxicity, or have the lowest level of toxicity possible for treating a patient having a disease or condition.

Other Uses For the Nucleic Acid Molecules of the Invention

The identification of different alleles of 5-LO can also be useful for identifying an individual among other individuals from the same species. For example, DNA sequences can be used as a fingerprint for detection of different individuals within the same species (Thompson, J. S. and Thompson, eds., Genetics in Medicine, WB Saunders Co., Philadelphia, PA (1991)). This is useful, for example, in forensic studies and paternity testing, as described below.

A. Forensics

Determination of which specific allele occupies a set of one or more polymorphic sites in an individual identifies a set of polymorphic forms that distinguish the individual from others in the population. *See generally* National Research Council, *The Evaluation of Forensic DNA Evidence* (Eds. Pollard *et al.*, National Academy Press, DC, 1996). The more polymorphic sites that are analyzed, the lower the probability that the set of polymorphic forms in one individual is the same as that in an unrelated individual. Preferably, if multiple sites are analyzed, the sites are unlinked. Thus, the polymorphisms of the invention can be used in conjunction with known polymorphisms in distal genes. Preferred polymorphisms for use in forensics are biallelic because the population frequencies of two polymorphic forms can usually be determined with greater accuracy than those of multiple polymorphic forms at multi-allelic loci.

The capacity to identify a distinguishing or unique set of polymorphic markers in an individual is useful for forensic analysis. For example, one can determine whether a blood sample from a suspect matches a blood or other tissue sample from a crime scene by determining whether the set of polymorphic forms occupying selected polymorphic sites is the same in the suspect and the sample. If the set of polymorphic markers does not match between a suspect and a sample, it can be concluded (barring experimental error) that the suspect was not the source of the sample. If the set of markers is the same in the sample as in the suspect, one can conclude that the DNA from the suspect is consistent with that found at the crime scene. If frequencies of the polymorphic forms at the loci tested have been determined (*e.g.*, by analysis of a suitable population of individuals), one can perform a statistical analysis to determine the probability that a match of suspect and crime scene sample would occur by chance.

p(ID) is the probability that two random individuals have the same polymorphic or allelic form at a given polymorphic site. For example, in biallelic loci, four genotypes are possible: AA, AB, BA, and BB. If alleles A and B occur in a haploid genome of the organism with frequencies x and y, the probability of each genotype in a diploid organism is (see WO 95/12607):

Homozygote: $p(AA)= x^2$
Homozygote: $p(BB)= y^2 = (1-x)^2$

Single Heterozygote: p(AB)= p(BA)= xy = x(1-x)
Both Heterozygotes: p(AB+BA)= 2xy = 2x(1-x)

The probability of identity at one locus (*i.e.,* the probability that two individuals, picked at random from a population will have identical polymorphic forms at a given locus) is given by the equation: $p(ID) = (x^2)$.

These calculations can be extended for any number of polymorphic forms at a given locus. For example, the probability of identity p(ID) for a 3-allele system where the alleles have the frequencies in the population of x, y, and z, respectively, is equal to the sum of the squares of the genotype frequencies: $P(ID) = x^4 + (2xy)^2 + (2yz)^2 + (2xz)^2 + z^4 + y^4$.

In a locus of n alleles, the appropriate binomial expansion is used to calculate p(ID) and p(exc).

The cumulative probability of identity (cum p(ID)) for each of multiple unlinked loci is determined by multiplying the probabilities provided by each locus:
cum p(ID) = p(ID1)p(ID2)p(ID3)...p(IDn).

The cumulative probability of non-identity for n loci (*i.e.,* the probability that two random individuals will be difference at 1 or more loci) is given by the equation:
cum p(nonID) = 1-cum p(ID).

If several polymorphic loci are tested, the cumulative probability of non-identity for random individuals becomes very high (*e.g.,* one billion to one). Such probabilities can be taken into account together with other evidence in determining the guilt or innocence of the suspect.

B. Paternity Testing

The object of paternity testing is usually to determine whether a male is the father of a child. In most cases, the mother of the child is known, and thus, it is possible to trace the mother's contribution to the child's genotype. Paternity testing investigates whether the part of the child's genotype not attributable to the mother is consistent to that of the puntative father. Paternity testing can be performed by analyzing sets of polymorphisms in the putative father and in the child.

If the set of polymorphisms in the child attributable to the father does not match the

set of polymorphisms of the putative father, it can be concluded, barring experimental error, that that putative father is not the real father. If the set of polymorphisms in the child attributable to the father does match the set of polymorphisms of the putative father, a statistical calculation can be performed to determine the probability of a coincidental match.

The probability of parentage exclusion (representing the probability that a random male will have a polymorphic form at a given polymorphic site that makes him incompatible as the father) is given by the equation (see WO 95/12607): p(exc) = xy(1-xy), where x and y are the population frequencies of alleles A and B of a biallelic polymorphic site.

(At a triallelic site p(exc) = xy(1-xy) + yz(1-yz) + xz(1-xz) + 3xyz(1-xyz)), where x, y, and z and the respective populations frequencies of alleles A, B, and C).

The probability of non-exclusion is: p(non-exc) = 1-p(exc).

The cumulative probability of non-exclusion (representing the values obtained when n loci are is used) is thus:

Cum p(non-exc) = p(non-exc1)p(non-exc2)p(non-exc3)...p(non-excn).

The cumulative probability of the exclusion for n loci (representing the probability that a random male will be excluded: cum p(exc) = 1 – cum p(non-exc).

If several polymorphic loci are included in the analysis, the cumulative probability of exclusion of a random male is very high. This probability can be taken into account in assessing the liability of a putative father whose polymorphic marker set matches the child's polymorphic marker set attributable to his or her father.

Kits

As set forth herein, the invention provides methods, *e.g.,* diagnostic and therapeutic methods, *e.g.,* for determining the type of allelic variant of a polymorphic region present in a 5-LO gene, such as a human 5-LO gene. In preferred embodiments, the methods use probes or primers comprising nucleotide sequences which are complementary polymorphic region of a 5-LO gene (SEQ ID NOs:4-6). In a more preferred embodiment, the methods are used to determine the identity of one or more allelic variants in polymorphic region which is included in the haplotype described herein. Accordingly, the invention provides kits for performing these methods.

In a preferred embodiment, the invention provides a kit for determining whether a patient has a more moderate or more severe inflammatory or allergy disease phenotype associated with a specific allelic variant of a 5-LO polymorphic region. In an even more preferred embodiment, the disease or disorder is characterized by an abnormal 5-LO activity, *e.g.,* aberrant 5-LO expression. In an even more preferred embodiment, the inflammatory or allergy disease is, *e.g.,* asthma, lung inflammation, nephritis, amyloidosis, rheumatoid arthritis, chronic bronchitis, sarcoidosis, scleroderma, lupus, polymyositis, Reiter's syndrome, psoriasis, pelvic inflammatory disease, orbital inflammatory disease, thrombotic disease, and inappropriate allergic responses to environmental stimuli such as poison ivy, pollen, insect stings and certain foods, including atopic dermatitis and contact dermatitis.

A preferred kit provides reagents for determining whether a patient will or will not be responsive to treatment of a disease or disorder associated with a specific allelic variant of a polymorphic region of a 5-LO gene, *e.g.,* asthma, with, for example, a 5-LO inhibitor. In a preferred embodiment, the inflammatory disease is, *e.g.,* asthma, lung inflammation, nephritis, amyloidosis, rheumatoid arthritis, chronic bronchitis, sarcoidosis, scleroderma, lupus, polymyositis, Reiter's syndrome, psoriasis, pelvic inflammatory disease, orbital inflammatory disease, thrombotic disease, and inappropriate allergic responses to environmental stimuli such as poison ivy, pollen, insect stings and certain foods, including atopic dermatitis and contact dermatitis. In a preferred embodiment, the kit of the invention can be used in selecting the appropriate drug to administer to a patient suffering from an inflammatory disease or disorder.

Preferred kits comprise at least one probe or primer which is capable of specifically hybridizing under stringent conditions to a 5-LO sequence or polymorphic region and instructions for use. The kits preferably comprise at least one of the above described nucleic acids. Preferred kits for amplifying at least a portion of a 5-LO gene, *e.g.,* the 5' upstream regulatory element, comprise two primers, at least one of which is capable of hybridizing to an allelic variant sequence. Even more preferred kits comprise a pair of primers selected from the group consisting of SEQ ID NO: 9 and SEQ ID NO: 10, SEQ ID NO: 11 and SEQ ID NO: 12, SEQ ID NO: 13 and SEQ ID NO: 14, SEQ ID NO: 15 and SEQ ID NO: 16 SEQ ID NO: 17 and SEQ ID NO: 18, SEQ ID NO: 19 and SEQ ID NO: 20, SEQ ID NO: 21 and

SEQ ID NO: 22, SEQ ID NO: 23 and SEQ ID NO: 24, SEQ ID NO: 25 and SEQ ID NO: 26, SEQ ID NO: 27 and SEQ ID NO: 28, SEQ ID NO: 29 and SEQ ID NO: 30, SEQ ID NO: 31 and SEQ ID NO: 32, SEQ ID NO: 33 and SEQ ID NO: 34, SEQ ID NO: 35 and SEQ ID NO: 36, SEQ ID NO: 37 and SEQ ID NO: 38, SEQ ID NO: 39 and SEQ ID NO: 40, SEQ ID NO: 41 and SEQ ID NO: 42, SEQ ID NO: 43 and SEQ ID NO: 44, SEQ ID NO: 45 and SEQ ID NO: 46, SEQ ID NO: 47 and SEQ ID NO: 48, SEQ ID NO: 49 and SEQ ID NO: 50, SEQ ID NO: 51 and SEQ ID NO: 52, SEQ ID NO: 53 and SEQ ID NO: 54, SEQ ID NO: 55 and SEQ ID NO: 56, SEQ ID NO: 57 and SEQ ID NO: 58, and SEQ ID NO: 59 and SEQ ID NO: 60 (see Table 3).

The kits of the invention can also comprise one or more control nucleic acids or reference nucleic acids, such as nucleic acids comprising a 5-LO intronic sequence. For example, a kit can comprise primers for amplifying a polymorphic region of a 5-LO gene and a control DNA corresponding to such an amplified DNA and having the nucleotide sequence of a specific allelic variant. Thus, direct comparison can be performed between the DNA amplified from a patient and the DNA having the nucleotide sequence of a specific allelic variant. In one embodiment, the control nucleic acid comprises at least a portion of a 5-LO gene of an individual who does not have an inflammatory disease, aberrant synthesis of leukotrienes, decreased eosinophil levels, or a disease or disorder associated with an aberrant 5-LO activity.

Yet other kits of the invention comprise at least one reagent necessary to perform the assay. For example, the kit can comprise an enzyme. Alternatively the kit can comprise a buffer or any other necessary reagent.

The present invention is further illustrated by the following examples which should not be construed as limiting in any way. The contents of all cited references (including, without limitation, literature references, issued patents, published patent applications as well as the Figures, Tables, and database references) as cited throughout this application are hereby expressly incorporated by reference. The practice of the present invention will employ, unless otherwise indicated, conventional techniques of cell biology, cell culture, molecular biology, transgenic biology, microbiology, recombinant DNA, and immunology, which are within the skill of the art. Such techniques are explained fully in the literature. See, for example, *Molecular Cloning A Laboratory Manual*, 2nd Ed., ed. by Sambrook,

Fritsch and Maniatis (Cold Spring Harbor Laboratory Press: 1989); *DNA Cloning*, Volumes I and II (D. N. Glover ed., 1985); *Oligonucleotide Synthesis* (M. J. Gait ed., 1984); Mullis *et al.* U.S. Patent Number 4,683,195; *Nucleic Acid Hybridization* (B. D. Hames & S. J. Higgins eds. 1984); *Transcription And Translation* (B. D. Hames & S. J. Higgins eds. 1984); *Culture Of Animal Cells* (R. I. Freshney, Alan R. Liss, Inc., 1987); *Immobilized Cells And Enzymes* (IRL Press, 1986); B. Perbal, *A Practical Guide To Molecular Cloning* (1984); the treatise, *Methods In Enzymology* (Academic Press, Inc., N.Y.); *Gene Transfer Vectors For Mammalian Cells* (J. H. Miller and M. P. Calos eds., 1987, Cold Spring Harbor Laboratory); *Methods In Enzymology*, Vols. 154 and 155 (Wu *et al.* eds.), *Immunochemical Methods In Cell And Molecular Biology* (Mayer and Walker, eds., Academic Press, London, 1987); *Handbook Of Experimental Immunology*, Volumes I-IV (D. M. Weir and C. C. Blackwell, eds., 1986); *Manipulating the Mouse Embryo*, (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., 1986).

EXAMPLES

Example 1: Determination of the genomic structure of the 5-LO gene

This example describes the determination of the genomic structure of the 5-LO gene. Identification of the various functional regions of the gene, including, but not limited to, 5' and 3' untranslated regions (UTR) and intron/exon boundaries was necessary for subsequent variant detection experiments. Two sequence comparison software applications were used to elucidate genomic structure of the 5-LO gene. Both applications involve the comparison of the cDNA sequence of the 5-LO gene to the sequence of large genomic DNA clones (bacterial artificial chromosomes or BACs) that encode the 5-LO gene in it's untranscribed form. One application is the Basic Local Alignment Search Tool™ or BLAST™ (Altschul *et al.* (1990) *J. Mol. Biol.* 215(3):403-410; Altschul *et al.* (1993) *Nature Genetics* 3:266-272; Altschul *et al.* (1997) *Nuc. Acids Res.* 25:3389-3402; Karlin and Altschul (1990) *Proc. Natl. Acad. Sci. USA* 87:2267-2268). The other application is Sequencher 3.1.1 (Gene Codes Corporation). Although based on different algorithms, these applications aid in the determination of genomic structure by the same principle.

These applications allow the comparison of two or more different nucleotide sequences (in the case of BLAST™, amino acid sequences can also be aligned) and

identification or alignment of regions of high similarity or identity. Aligning cDNA sequence of the 5-LO gene with the corresponding BAC sequences with either tool allows visualization of regions of identity between the transcribed sequence and the sequence of the genomic region from which it was derived. Regions of sequence identity between the two sources are exonic regions and can be confirmed by the presence of known conserved splice elements flanking them. Regions of sequence contained in the BAC sequences but not in the cDNA sequence are considered to be either intronic sequence or lay outside of the 5' and 3' boundaries of the gene.

Once the genomic structure of the 5-LO gene, *e.g.*, intron/exon boundaries, was determined, PCR primers were designed for use with variant nucleotide detection experiments. Exonic sequences (including intron/exon junctions), 5'UTR, 3' UTR and approximately of 1 kbp upstream of the transcription start site were scanned for variant nucleotide detection as described in Examples 2, 3, 4, and 5.

Example 2: Identification of primer pairs to isolate intronic, exonic, and 5' upstream regulatory element sequences for detection of polymorphisms and mutations

Multiple pairs of primers were synthesized in order to amplify each of the exons or portions thereof and adjacent intronic regions. Genomic DNA from a human subject was subjected to PCR in 25 µl reactions (1x PCR Amplitaq polymerase buffer, 0.1 mM dNTPs, 0.8 µM 5' primer, 0.8 µM 3' primer, 0.75 units of Amplitaq polymerase, 50 ng genomic DNA) using each of the above described pairs of primers under the following cycle conditions: 94°C for 2 min, 35 x [94°C for 40 sec, 57°C for 30 sec, 72°C for 1 min], 72°C for 5 min, 4°C hold. The resulting PCR products were analyzed on a 2% agarose gel. The identity of the PCR product was confirmed by digestion with a restriction enzyme and subsequent agarose electrophoresis. Twenty-three pairs of oligomers were chosen to serve as PCR primers to amplify regions containing each of the 14 coding exons of the human 5-LO gene and eleven pairs of primers were chosen to serve as PCR primers to amplify the 5' upstream regulatory element. The nucleotide sequences of the forward and reverse primers are indicated in Table 3, as well as the SEQ ID NOs for each primer and the location of the primers (*e.g.*, exon, 3'UTR, or 5'UTR/5' upstream regulatory element) within the 5-LO gene. The expected sizes of the PCR products are also set forth in Table 3. In addition,

polymorphism ID numbers corresponding to those listed in Table 1 are included where there was a polymorphism identified with a particular primer pair.

Example 3: Detection of polymorphic regions in the human 5-LO gene by DHPLC

This example describes the use of denaturing high performance liquid chromatography (DHPLC) for the identification of DNA sequence variations.

DNA samples for these experiments were obtained from a population of 144 individuals to be used for polymorphism discovery. These 144 DNA samples included DNA samples from a population of forty-eight North American Caucasian individuals, forty-eight African American individuals, and forty-eight individuals from throughout the Anhui province in East Central China. Furthermore, the forty-eight DNA samples from the Chinese individuals are from unrelated control subjects, obtained at random, and have been assessed for a number of traits related to asthma and allergies. Among the primary variables of interest are lung function , a response to "skin-prick" tests, physician's diagnosis of asthma, total serum IgE, and peripheral blood eosinophils. The peripheral blood eosinophil count was measured using the Coulter counter technique as described in Barnard DF, *et al.* (1989) *Clin Lab Haematol* 11(3):255-66, and is expressed in terms of eosinophils per microliter.

DHPLC uses reverse-phase ion-pairing chromatography to detect the heteroduplexes that are generated during amplification of PCR fragments from individuals who are heterozygous at a particular nucleotide locus within that fragment (Oefner and Underhill (1995) *Am. J. Human Gen.* 57:Suppl. A266).

Generally, the analysis was carried out as described in O'Donovan *et al.* ((1998) Genomics 52:44-49). PCR products having product sizes ranging from about 150-400 bp were generated using the primers and PCR conditions described in Example 2. Two PCR reactions were pooled together for DHPLC analysis (4 ul of each reaction for a total of 8 ul per sample). DHPLC was performed on a DHPLC system purchased from Transgenomic, Inc. The gradient was created by mixing buffers A (0.1M TEAA) and B (0.1M TEAA, 25% Acetontitrile). WAVEmaker™ software was utilized to predict a melting temperature and calculate a buffer gradient for mutation analysis of a given DNA sequence. The resulting chromatograms were analyzed to identify base pair alterations or deletions based on specific chromatographic profiles.

Example 4: Detection of polymorphic regions in the human 5-LO gene by SSCP

Genomic DNA from each of the 144 individuals described in Example 3 was subjected to PCR in 25 µl reactions (1X PCR Amplitaq polymerase buffer, 0.1 mM dNTPs, 0.8 µM 5' primer, 0.8 µM 3' primer, 0.75 units of Amplitaq polymerase, 50 ng genomic DNA) using each of the above described pairs of primers under the following cycle conditions: 94°C for 2 min, 35 x [94°C for 40 sec, 57°C for 30 sec, 72°C for 1 min], 72°C 5 min, 4°C hold. The expected sizes of the PCR products, are also indicated in Table 3.

The amplified genomic DNA fragments were then analyzed by SSCP (Orita *et al.* (1989) *PNAS USA* 86:2766, see also Cotton (1993) *Mutat Res* 285:125-144; and Hayashi (1992) *Genet Anal Tech Appl* 9:73-79). From each 25 µl PCR reaction, 3 µl was taken and added to 7 µl of loading buffer. The mixture was heated to 94°C for 5 min and then immediately cooled in a slurry of ice-water. 3-4 µl were then loaded on a 10% polyacrylamide gel either with 10% glycerol or without 10% glycerol, and then subjected to electrophoresis either overnight at 4 Watts at room temperature, overnight at 4 Watts at 4°C (for amplifying a 5' upstream regulatory element), or for 5 hours at 20 Watts at 4°C. The secondary structure of single-stranded nucleic acids varies according to sequence, thus allowing the detection of small differences in nucleic acid sequence between similar nucleic acids. At the end of the electrophoretic period, the DNA was analyzed by gently overlaying a mixture of dyes onto the gel (1x the manufacturer's recommended concentration of SYBR Green I™ and SYBR Green II™ in 0.5 X TBE buffer (Molecular Probes™)) for 5 min, followed by rinsing in distilled water and detection in a Fluoroimager 575™ (Molecular Dynamics™).

Example 5: Identification of polymorphic regions in the human 5-LO gene by direct sequencing of PCR products

To determine the sequences of the polymorphisms identified, the regions containing the polymorphisms were reamplified using the aforementioned primers. The genomic DNA from the subjects was subjected to PCR in 50 µl reactions (1x PCR Amplitaq polymerase buffer, 0.1 mM dNTPs, 0.8 µM 5' primer, 0.8 µM 3' primer, 0.75 units of Amplitaq polymerase, 50 ng genomic DNA) using each of the above described pairs of primers under

the following cycle conditions: 94°C for 2 min, 35 x [94°C for 40 sec, 57°C for 30 sec, 72°C for 1 min], 72°C 5 min, 4°C hold. The newly amplified products were then purified using the Qiagen Qiaquick™ PCR purification kit according to the manufacturer's protocol, and subjected to sequencing using the aforementioned primers which were utilized for amplification.

The results indicate that there is a change from a guanine to an adenine in the 5' upstream regulatory element of the 5-LO gene at residue 1000 of the reference sequence GI 187166 (polymorphism ID No. 5loprr1). A second polymorphism is a deletion in the 5' upstream regulatory element at residues 472-477 of the reference sequence GI 187166 (polymorphism ID No. 5lo01a). A third polymorphism is a change from a guanine to an adenine in the 5' upstream regulatory element at residue 559 of the reference sequence GI 187166 (polymorphism ID No. 5lo04a). These polymorphisms are listed in Table 1 and correspond to SEQ ID NOs: 4-6.

Example 6: Identification of a Variant Haplotype in the Promoter Region of 5-LO

Three populations comprised of individuals of different racial backgrounds, *i.e.,* North American Caucasian, Asian Chinese, and African American, were screened for variants in the promoter region of 5-LO using the methods described herein. It was found that polymorphism Id Nos. 5loprr1, 5lo01a, 5lo04a, 5lonrra, and 5lonrrb all segregate together such that when an individual has one of these polymorphisms, all five polymorphisms are present. This five polymorphism haplotype is conserved at an identical frequency between these three different racial backgrounds (North American Caucasian, Asian Chinese, and African American). The five polymorphisms (polymorphism Id Nos. 5loprr1, 5lo01a, 5lo04a, 5lonrra, and 5lonrrb) are in complete linkage disequilibrium and are present at a frequency of approximately 18% in all three populations screened. Moreover the haplotype was in linkage disequilibrium with any one of the three variant 5-LO Sp-1 binding sites. In Caucasian and Chinese populations, the haplotype was associated with the variant Sp1 binding site having four Sp1 motifs. In the African American population, the haplotype was found to be associates with any one of the three variant Sp1 binding sites (i.e., biinding sites having 3, 4 or 6 Sp1 motifs).

Example 7: Association of 5-LO variants with eosinophil levels in Chinese asthmatic families

This example describes the determination of the association between the presence of one or more polymorphisms, low eosinophil levels in a sample of Chinese asthmatic families.

Method

A study of 275 families was conducted to determine whether the variant identified as "Polymorphism ID No. 5lo01a" (see Table 1) is associated with asthma-related phenotypes. Selected individuals from these families were genotyped at the 5lo01a polymorphism. Two statistical tests were then conducted to determine whether evidence exists of an association between this locus and asthma-related phenotypes.

For this study, genomic DNA was isolated from peripheral blood lymphocytes using the Puregene™ kit (Gentra Systems Inc.™), according to manufacturer's suggested protocol. SSCP analysis (see Example 4) was used to determine each subjects' genotype at the promoter deletion/insertion polymorphism 5lo01a (*e.g.*, the deletion/insertion of (GTTAAA) (SEQ ID NO:62).

To determine whether genotype is associated with peripheral blood eosinophil count (an index of asthmatic disease), counts (per ml) were compared between individuals carrying at least one copy of the deletion polymorphism (*e.g.*, those with at least one allele that is missing the residues "GTTAAA" (SEQ ID NO:62)), and those carrying no copies of the polymorphism. One individual was selected from each family; the selections were based solely on genotype, and were designed to make the sizes of these two comparison groups as similar as possible. A Student T-TEST was used to determine whether the difference in mean eosinophil count between these two groups was significant.

Results

For the first test, 138 unrelated individuals with at least one copy of the deletion were analyzed, as were 137 unrelated individuals with no copies of the deletion. The mean

eosinophil level among individuals with the deletion was 222.17 (standard deviation=246.38); the mean among individuals without the deletion was 324.09 (standard deviation=482.75). The Student T-TEST comparing the mean eosinophil level between these two groups revealed that there is a significant difference (p=.029) between these means.

Discussion

An association test was performed to determine whether a 5-LO promoter variant is associated with asthma-related phenotypes. Specifically, it was examined whether genotype at the 5-LO polymorphism associates with peripheral blood eosinophil counts. The test gave evidence that such an association exists. The test demonstrated that eosinophil counts are significantly lower among individuals carrying at least one copy of the rare deletion allele at this locus. Thus, the promoter genotype can be used predict whether an individual has a greater chance of having a abnormally low eosinophil count.

Equivalents

Those skilled in the art will recognize, or be able to ascertain using no more than routine experimentation, many equivalents to the specific embodiments of the invention described herein. Such equivalents are intended to be encompassed by the following claims.